

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 18, 2001, 03:46:36 ; Search time 1381.47 Seconds
(without alignments)
4060.207 Million cell updates/sec

Title: us-09-402-488a-1

Perfect score: 1096
Sequence: 1 ATGTCCCTACTAGCTTA.....ATCCACATCTTAGACCG 1096

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1118133 seqs, 2558875100 residues

Total number of hits satisfying chosen parameters: 2236266

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba1: *
2: gb_ba2: *
3: gb_om: *
4: gb_ov: *
5: gb_ph: *
6: gb_pl1: *
7: gb_pl2: *
8: gb_pl3: *
9: gb_pr1: *
10: gb_pr2: *
11: gb_pr3: *
12: gb_sy: *
13: gb_un: *
14: em_fun: *
15: em_hum1: *
16: em_hum2: *
17: em_in: *
18: em_om: *
19: em_or: *
20: em_ov: *
21: em_pat: *
22: em_ph: *
23: em_pl: *
24: em_ro: *
25: em_sts: *
26: em_sy: *
27: em_un: *
28: em_vl: *
29: gb_ba3: *
30: gb_in1: *
31: gb_in2: *
32: gb_in3: *
33: gb_pl3: *
34: gb_pr4: *
35: em_ba1: *
36: em_ba2: *
37: em_hc91: *
38: em_hc92: *
39: em_hc93: *
40: em_hc94: *
41: em_hc95: *
42: em_hc96: *
43: em_hc97: *

44: em_hc98: *
45: em_hc99: *
46: em_hc10: *
47: em_hum3: *
48: em_hum4: *
49: em_hum5: *
50: em_hum6: *
51: gb_pr5: *
52: gb_pr6: *
53: gb_pr7: *
54: gb_hc91: *
55: gb_hc92: *
56: gb_hc93: *
57: gb_hc94: *
58: gb_hc95: *
59: gb_hc96: *
60: gb_hc97: *
61: gb_hc98: *
62: gb_hc99: *
63: gb_hc10: *
64: gb_hc11: *
65: gb_hc12: *
66: gb_hc13: *
67: gb_hc14: *
68: gb_hc15: *
69: gb_hc16: *
70: gb_hc17: *
71: gb_hc18: *
72: gb_hc19: *
73: gb_hc20: *
74: gb_hc21: *
75: gb_hc22: *
76: gb_hc23: *
77: gb_hc24: *
78: gb_hc25: *
79: gb_hc26: *
80: gb_hc27: *
81: gb_hc28: *
82: gb_hc29: *
83: em_hc90: *
84: gb_hc24: *
85: gb_hc28: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query	Length	DB ID	Description
1	708.6	64.7	4968	12 XXU13855	U13855 pGEX-4T3 c1
2	697.6	63.6	4969	12 XXU13853	U13853 pGEX-4T1 c1
3	696.6	63.6	4970	12 XXU13854	U13854 pGEX-4T2 c1
4	685	62.5	4972	12 XXU13856	U13856 pGEX-5X1 c1
5	682.4	62.3	4947	12 XXU13849	U13849 pGEX-1 Lamb
6	681	62.1	3417	12 AB014641	AB014641 Cloning v
7	681	62.1	4948	12 XXU13850	U13850 pGEX-2T c10
8	681	62.1	5042	12 EVU084572	U084572 Expression
9	681	62.1	5052	12 EVU084571	U084571 Expression
10	681	62.1	5501	12 EVU085201	U085201 Expression
11	681	62.1	5504	12 EVU085202	U085202 Expression
12	681	62.1	5516	12 EVU085203	U085203 Expression
13	681	62.1	5738	12 EVU085204	U085204 Expression
14	681	62.1	5756	12 EVU085205	U085205 Expression
15	681	62.1	5990	12 EVU085206	U085206 Expression
16	679.2	62.0	4973	12 XXU13857	U13857 pGEX-5X2 c1
17	678.4	61.9	1281	81 A47307	A47307 Sequence 2
18	678.2	61.9	1740	81 AR067794	AR067794 Sequence
19	678	61.9	9760	81 EVU086815	U086815 Expression
20	678	61.9	9763	12 AF126280	AF126280 Expression
21	677.4	61.8	4969	12 XXU13851	U13851 pGEX-2TK c1

```

22 677 61.8 1119 81 E13745
23 677 61.8 10765 13 EVU67875
24 676.2 61.7 7805 32 YSCCJ2INDU
25 675.6 61.6 4983 12 CVU78874
26 675 61.6 4974 12 XXU13858
27 674.6 61.6 3378 12 AF149265
28 673.8 61.5 3966 12 EVU81274
29 668.2 61.0 4984 12 CVU78872
30 667.4 60.9 897 81 E16870
31 664.2 60.6 897 21 E09612
32 664 60.6 4985 12 CVU78873
33 663.8 60.6 5487 12 PGH223813
34 663.4 60.5 1128 21 E09265
35 663.4 60.5 4952 12 XXU13852
36 662.8 60.5 2095 81 AR026350
37 661.4 60.3 3950 12 EVU81273
38 661 60.3 4932 12 SYNECOPGX
39 660.2 60.2 3970 12 EVU81275
40 658.6 60.1 4905 12 ASPGST
41 655.2 59.8 5037 12 AF097411
42 655 59.8 739 32 SCMAg
43 654 59.7 654 81 AR036684
44 652.2 59.5 897 21 E09610
45 651 59.4 9750 12 AF025397

```

ALIGNMENTS

```

source 1..4968
        /organism="unidentified cloning vector"
        /db_xref="taxon:45196"
        /lab_host="Escherichia coli"
        /standard_name="lac"
        /note="lac promoter for inducible expression of
        glutathione S-transferase"
        217..237
        /bound_molecule="lac repressor protein"
        258..980
        /citation=[3]
        /citation=[4]
        /codon_start=1
        /transl_table=11
        /product="glutathione S-transferase"
        /protein_id="AA57104.1"
        /db_xref="GI:595730"
        /translation="MSPILGKWKIKGLVOPRLLELYEKEEELERDEGDKRNK
        KFELEFPNLPYIIGDVKLTOSMAIIRYADKHNMLGCPKRRARISMLEGVLDI
        VLVGSRISAKDEFTLAKVDPLSKIPKAKHEDELCRTYINQDVTHTPDMLDALD
        VLVMDPMDCAFKEVCFKRIATPQIDKYSKYTIAMPLOGWQATRGCGDHPK
        SDLVKGSPPNSRVDSGRIVTD"
        918..935
        /note="encodes thrombin recognition site"
        930..965
        /note="Multiple Cloning Site (MCS); contains the unique
        restriction sites BamHI, EcoRI, SmaI, SalI, XhoI, and
        NotI"
        1306..2236
        /gene="bla"
        /gene="bla"
        /gene="bla"
        1376..2236
        /gene="bla"
        /transl_table=1
        /codon_start=1
        /product="beta-lactamase"
        /protein_id="AA57105.1"
        /db_xref="GI:595731"
        /translation="MSTQHRVALIPFAECLEVPAPHETLVKVKDAEDOLGARVY
        IELDLSNGKLTIESRPERPMPSTFVLGCAVLSVDAGOEOLGRIRIHYSONDLVE
        YSPTEKHLIDGMTVRELCSAITSNTANILLTIGGPKELTARHMGHVPRL
        DRPEPELNEAIPDERDPTTAPAMATLRLTIGELLTALROOIIDMWADVAGTL
        LRSALPGMGFIAKSGSGERGSGCIITALADPDKPSRIYVITTTGSOATDENRQIA
        ETGASLIRKM"
        2994
        /note="base 2994 represents the first base of the newly
        synthesized single strand"
        3317..4399
        /direction=right
        3317..4399
        /gene="lacIq"
        3317..4399
        /gene="lacIq"
        /transl_table=11
        /codon_start=1
        /product="lac repressor"
        /protein_id="AA57106.1"
        /db_xref="GI:595732"
        /translation="MKPVTLDAEYAGVSYQTVSRVNOASHVSAKTRKVEAAMAE
        LNYIPNVAOOLAGKOSILIGVATSSALAPSOIYAATKSRADOGASVVSVMERS
        GVACAKAAVNNLAAORVSGLIINYPLDDDAITVAEACTVPALFLDVSQPTINSIT
        FSEHEDGRICVEHLVALGHOAILLAKGPPSSVSARILRAGMHTYLTIRNOIPIAEBEG
        DMSAGSFQDTOMLNGEIVPTAMLVANDMAIGAMRATTESSGLRVGADISVVGYYDT
        EDSSCYIPLTTTITKQDFRLIGQTSVDRLQLDSOGAVKGNQLPVSILVKKRTYLAANT
        QRTSPALADLSLMQAROVSRLESQ"
        1202 c 1291 g 1250 t

```

RESULT 1

XXU13855

LOCUS XXU13855 4968 bp DNA circular SYN 13-DEC-1994
 DEFINITION pgEX-473 cloning vector, complete sequence.
 ACCESSION U13855
 VERSION U13855.1 GI:595729

KEYWORDS glutathione S-transferase; beta-lactamase; lac repressor.
 SOURCE unidentified cloning vector.
 ORGANISM artificial sequence; vectors.

REFERENCE 1 (bases 1 to 4968)
 AUTHOR(S) Malone,J.A.

JOURNAL pgEX-473: A cloning vector for the inducible expression of genes as
 REFERENCE glutathione S-transferase fusion proteins containing a thrombin
 AUTHOR(S) unpublished (1994)
 TITLE 2 (bases 1 to 4968)

JOURNAL Smith,D.B. and Johnson,K.S.
 REFERENCE Single-step purification of polypeptides expressed in Escherichia
 AUTHOR(S) coli as fusions with glutathione S-transferase
 TITLE Gene 67 (1), 31-40 (1988)

JOURNAL 88329742
 MEDLINE 3 (bases 264 to 911)
 REFERENCE Smith,D.B., Davern,K.M., Board,P.G., Tiu,W.U., Garcia,E.G. and
 AUTHOR(S) Mitchell,G.F.

JOURNAL Mr 26,000 antigen of Schistosoma japonicum recognized by resistant
 REFERENCE WEHI 129/J mice is a parasite glutathione S-transferase [published
 MEDLINE erratum appears in Proc Natl Acad Sci U S A 1987 Sep;84(18):6541]
 AUTHOR(S) Proc. Natl. Acad. Sci. U.S.A. 83 (22), 8703-8707 (1986)
 TITLE 4 (bases 881 to 911)

JOURNAL Smith,D.B., Davern,K.M., Board,P.G., Tiu,W.U., Garcia,E.G. and
 REFERENCE Mitchell,G.F.
 MEDLINE Correction: Mr 26,000 antigen of Schistosoma japonicum recognized
 AUTHOR(S) by resistant WEHI 129/J mice is a parasite glutathione
 TITLE S-transferase
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 84, 6541-6541 (1987)
 REFERENCE 5 (bases 1 to 4968)

JOURNAL Direct Submission
 REFERENCE Submitted (13-AUG-1994) James A. Malone, International Technical
 AUTHOR(S) Services, Molecular Biology Reagents Division, Pharmacia Biotech
 JOURNAL Inc., 2202 N. Bartlett Ave., Milwaukee, WI 53202-1009, USA
 FEATURES Location/Qualifiers

BASE COUNT 1225 a 1291 g 1250 t
 ORIGIN

Query Match 64.7%; Score 708.6; DB 12; Length 4968;
 Best Local Similarity 99.4%; Pred. No. 3,7e-172;
 Matches 711; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

Oy 1 ANGTCCTCATACTAGTATTGAAATAAGGGCCCTTGCAACCCACTGCACCTCTT 60
    |||
Db 258 ANGTCCTCATACTAGTATTGAAATAAGGGCCCTTGCAACCCACTGCACCTCTT 317
Oy 61 TTGGCAATATCTTGAAGAAAAATATGACAGCATTTGTATGACGGCGATGAGGTATAA 120
    |||
Db 318 TTGGCAATATCTTGAAGAAAAATATGACAGCATTTGTATGACGGCGATGAGGTATAA 377
Oy 121 TGGCGAACAATAAGTTGATTTGGTTGGAGTTTCCCATCTTCTTATATATGAT 180
    |||
Db 378 TGGCGAACAATAAGTTGATTTGGTTGGAGTTTCCCATCTTCTTATATATGAT 437
Oy 181 GGTGATGTTAAATTAACACAGTCTATGACCATCATACGTTATATAGCTGACAGCAAC 240
    |||
Db 438 GGTGATGTTAAATTAACACAGTCTATGACCATCATACGTTATATAGCTGACAGCAAC 497
Oy 241 AGTTGGCGGTGTTGCCAAAGAGCGTGCAGAGATTTCAATGCTTGAAGAGCGGTTTG 300
    |||
Db 498 AGTTGGCGGTGTTGCCAAAGAGCGTGCAGAGATTTCAATGCTTGAAGAGCGGTTTG 557
Oy 301 GATATTAGATACGGTGTCTTGCAGAAATGCATATAGTAAGACTTTGAACCTGCAAGTT 360
    |||
Db 558 GATATTAGATACGGTGTCTTGCAGAAATGCATATAGTAAGACTTTGAACCTGCAAGTT 617
Oy 361 GATTTTCTTACAGACCTACCTGAATGCTGAATAAGTTGAGAGATCGTTATGTCATAA 420
    |||
Db 618 GATTTTCTTACAGACCTACCTGAATGCTGAATAAGTTGAGAGATCGTTATGTCATAA 677
Oy 421 ACATATTTAATGCGATCATGTAACCCATCTGACTTATGTTATGACGCTTGAT 480
    |||
Db 678 ACATATTTAATGCGATCATGTAACCCATCTGACTTATGTTATGACGCTTGAT 737
Oy 481 GTTGTATTACATGAGACCAATGCGCTGATGCGTTCGCAAAATAGTTGTTTAA 540
    |||
Db 738 GTTGTATTACATGAGACCAATGCGCTGATGCGTTCGCAAAATAGTTGTTTAA 797
Oy 541 AAACGATTTGAAGCTATCCCAAAATGTAAGTACTTGAATCCAGCAATATATAGCA 600
    |||
Db 798 AAACGATTTGAAGCTATCCCAAAATGTAAGTACTTGAATCCAGCAATATATAGCA 857
Oy 601 TGGCCCTTTCAGAGCGTGCAGACGCTTGTGTGGCGGACCATCTCCAAATCGCAT 660
    |||
Db 858 TGGCCCTTTCAGAGCGTGCAGACGCTTGTGTGGCGGACCATCTCCAAATCGCAT 917
Oy 661 CTGTTCCGCGTGAATCCCGGAATCCGCGGTGCACTCGAGGCGCGCTGGA 715
    |||
Db 918 CTGTTCCGCGTGAATCCCGGAATCCGCGGTGCACTCGAGGCGCGCATTCGGA 972

RESULT 2
XXU13853 4969 bp DNA circular SYN 13-DEC-1994
LOCUS XXU13853
DEFINITION pGEX-4T1 cloning vector, complete sequence.
ACCESSION U13853
VERSION U13853.1 GI:595721
KEYWORDS glutathione S-transferase; beta-lactamase; lac repressor.
SOURCE unidentified cloning vector.
ORGANISM artificial sequence; vectors.
REFERENCE 1 (bases 1 to 4969)
AUTHORS Malone,J.A.
TITLE pGEX-4T1: A cloning vector for the inducible expression of genes as
glutathione S-transferase fusion proteins containing a thrombin
cleavage site
JOURNAL Unpublished (1994)
REFERENCE 2 (bases 1 to 4969)
AUTHORS Smith,D.B. and Johnson,K.S.
TITLE Single-step purification of polypeptides expressed in Escherichia
coli as fusions with glutathione S-transferase
JOURNAL Gene 67 (1), 31-40 (1988)
MEDLINE 88329742
REFERENCE 3 (bases 264 to 911)

```

```

AUTHORS Smith,D.B., Davern,K.M., Board,P.G., Tin,W.U., Garcia,E.G. and
TITLE Mitchell,G.F.
JOURNAL WEHI 129/J mice is a parasite glutathione S-transferase [published
erratum appears in Proc Natl Acad Sci U S A 1987 Sep;84(18):5541]
MEDLINE 87041520
REFERENCE 4 (bases 881 to 911)
AUTHORS Smith,D.B., Davern,K.M., Board,P.G., Tin,W.U., Garcia,E.G. and
TITLE Mitchell,G.F.
JOURNAL Correction: Mr 26,000 antigen of Schistosoma japonicum recognized
by resistant WEHI 129/J mice is a parasite glutathione
S-transferase
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 84, 6541-6541 (1987)
REFERENCE 5 (bases 1 to 4969)
AUTHORS Malone,J.A.
TITLE Direct Submission
JOURNAL Submitted (19-AUG-1994) James A. Malone, International Technical
Services, Molecular Biology Reagents Division, Pharmacia Biotech
Inc., 2202 N. Bartlett Ave., Milwaukee, WI 53202-1009, USA

FEATURES
source 1..4969
    /organism="unidentified cloning vector"
    /db_xref="taxon:45196"
    /lab_host="Escherichia coli"
    /std_name="tac"
    /note="tac promoter for inducible expression of
    glutathione S-transferase"
    /bound_moiety="lac repressor protein"
    /protein_bind
promoter 183..211
    /standard_name="tac"
    /note="tac promoter for inducible expression of
    glutathione S-transferase"
    /bound_moiety="lac repressor protein"
CDS 217..237
    /citation={3}
    /codon_start=1
    /transl_table=11
    /product="glutathione S-transferase"
    /protein_id="AA57098.1"
    /db_xref="GI:595722"
    /translation="MSPIIGYKIRGIVOPTLLLEYLEKEEHEHLYEDEDDKRNK
    KTELLEPPNLPYIDGVKLTOSMAITRIYADKINIMGGCKREAREISMECAVLDI
    RYGVRIAYSKDEPLIKVDELKPEMLKMFEDRLCHKTYLNGDHTHPDPMVLALD
    VLVMDPMCLDAFPKLVCFKRIEAIPOIDKYLKSSKYIAMPLOQMOTAFGGDHPPK
    SDLVPRGSPREFGRLEPRHRD"
    /note=".935
    /note="encodes thrombin recognition site"
    /note="Multiple Cloning Site (MCS): contains the unique
    restriction sites BamHI, EcoRI, SmaI, SalI, XhoI, and
    NotI; The EcoRI site of pGEX-4T1 is in frame with the
    EcoRI site of lambda gtl1"
    /gene="bla"
    /gene="bla"
    /gene="bla"
    /codon_start=1
    /transl_table=11
    /product="beta-lactamase"
    /protein_id="AA57099.1"
    /db_xref="GI:595723"
    /translation="MSIOHFRVALIPFAAFLPYFAHPETLVKVKDAEDOLGAVGY
    IELIDNSKILESREPEREPKMSFKVILCAVLSRDADGEOLGRIRHYSQNDLVE
    YSPTEKHLDITGVRELCSAATISNDITAAWLITTTGGRELTAFPHNMDDHTRL
    DRWPELNEAIPNDRITTPAAMAATTLRLKLLTGLTSLASQQLIDWNEADKVAGPL
    LRSALPACMFADKSGAGERSGRTIIALGPDGKPSRIIVITVTSQATMDERNROIA
    ELGASLIKHW"
    /note="base 2995 represents the first base of the newly
    synthesized single strand"
    /direction=right
    /gene
rep_origin 2995
    /note="base 2995 represents the first base of the newly
    synthesized single strand"
    /direction=right
    /gene

```

```

CDS
      /gene="laci4"
      /protein_id="AA057100.1"
      /db_xref="GI:595724"
      /translation="MKPYLYDVAEVAGSVQTSRVVYNQASHVSAKTRVEAMAE
      INYIPNVAOOLAGOSLLIGVATSSLLAHPSOIVAAIKSRADOLGASVSWERS
      GVEACKAAVHNLAORVGLIINYPDODAIAYAACTNVPALFDVSDTPINSII
      FSHEDGTBLGVHLYALHQOJAILLAGFLSSVSARLRLAGMKYLTINQIPIEREG
      DMSASGFQOYMOLNEGIIVPTAMVANDOMALGMRAITTESGLVGDIVSGYDDT
      EDSSCIPEPLTTIKQDFRLQOTSVDRLOLSOGQAVYKGNLLEPVSIVKRRTIAPNT
      QVAPRALADSIQMLARQVSRLESGO"
BASE COUNT      1225 a      1202 c      1292 g      1250 t
ORIGIN

```

```

Query Match      63.6%; Score 697.6; DB 12; Length 4969;
Best Local Similarity 99.3%; Pred. No. 2.5e-169;
Matches 711; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 1  ATGTCCCATATCTAGGTATGTGAATAATTAGGGCTTGTGCAACCCACTGACTTCT 60
    |||
Db 258 ATGTCCCATATCTAGGTATGTGAATAATTAGGGCTTGTGCAACCCACTGACTTCT 317
    |||

QY 61  TTGGAATATCTTGAAGAAAAATATGAGAGCATTTGTATGAGCGGATGAAGTATAA 120
    |||
Db 318 TTGGAATATCTTGAAGAAAAATATGAGAGCATTTGTATGAGCGGATGAAGTATAA 377
    |||

QY 121 TGGCAAAACAAAAGTTGGAATGGGTTTGGATTTCCCATCTTCTTATTAATGAT 180
    |||
Db 378 TGGCAAAACAAAAGTTGGAATGGGTTTGGATTTCCCATCTTCTTATTAATGAT 437
    |||

QY 181 GGTGATGTTAATTAACACAGCTCTATGGCCATCATAGCTTATATGAGCAAGACAC 240
    |||
Db 438 GGTGATGTTAATTAACACAGCTCTATGGCCATCATAGCTTATATGAGCAAGACAC 497
    |||

QY 241 ATGTGGGTGCTGTGCAAAAAGCGCTGACAGATTTCAATGCTTGAAGAGCGGTTTGG 300
    |||
Db 498 ATGTGGGTGCTGTGCAAAAAGCGCTGACAGATTTCAATGCTTGAAGAGCGGTTTGG 557
    |||

QY 301 GATATTAGATCGGTGTGCGAATTCGATATATGTAAGACTTTGAACTCTCAAGTT 360
    |||
Db 558 GATATTAGATCGGTGTGTGCGAATTCGATATGTAAGACTTTGAACTCTCAAGTT 617
    |||

QY 361 GATTTTCTTACCAAGCTAAGCTGAATGCTGAATAATGTTGCAAGATCGTTTATGTCATAA 420
    |||
Db 618 GATTTTCTTACCAAGCTAAGCTGAATGCTGAATAATGTTGCAAGATCGTTTATGTCATAA 677
    |||

QY 421 ACATATTAATGTGTGATCATGTAAACCATCTGACCTTCATGTTGTATGATGACGCTCTGAT 480
    |||
Db 678 ACATATTAATGTGTGATCATGTAAACCATCTGACCTTCATGTTGTATGATGACGCTCTGAT 737
    |||

QY 481 GTTGTATTAACATGAGCAACCAATGTGCTGTGATCGTTCCTCAAAATAGTTGTTTAA 540
    |||
Db 738 GTTGTATTAACATGAGCAACCAATGTGCTGTGATCGTTCCTCAAAATAGTTGTTTAA 797
    |||

QY 541 AAACGATATGAGCTATCCCAACAATGATAGTACTTGAATTCAGCAAGATATATAGCA 600
    |||
Db 798 AAACGATATGAGCTATCCCAACAATGATAGTACTTGAATTCAGCAAGATATATAGCA 857
    |||

QY 601 TGGGCTTTGACAGGGCTGGCAAGCGTTGTGTGTGGTGGCAACCATCTCTCAAAATCGGAT 660
    |||
Db 858 TGGGCTTTGACAGGGCTGGCAAGCGTTGTGTGTGGTGGCAACCATCTCTCAAAATCGGAT 917
    |||

QY 661 CTGGTTCCGGGTGATCCCC-GAATTCGCGGTGACATGACGCGCGCGCTGAGA 715
    |||
Db 918 CTGGTTCCGGGTGATCCCCGGAATTCGCGGTGACATGACGCGCGCGCTGAGA 973
    |||

```

RESULT 3

```

XXU13854
LOCUS      XXU13854      4970 bp      DNA      circular      SYN
DEFINITION pgEX-472 cloning vector, complete sequence.
ACCESSION  U13854
VERSION    U13854.1
KEYWORDS   GI:595725
SOURCE     glutathione S-transferase; beta-lactamase; lac repressor.
ORGANISM   unidentified cloning vector.
REFERENCE  1 (bases 1 to 4970)
AUTHORS    Malone,J.A.
TITLE      pGEX-472: A cloning vector for the inducible expression of genes as
           glutathione S-transferase fusion proteins containing a thrombin
           cleavage site
REFERENCE  2 (bases 1 to 4970)
AUTHORS    Smith,D.B., and Johnson,K.S.
TITLE      Single-step purification of polypeptides expressed in Escherichia
           coli as fusions with glutathione S-transferase
REFERENCE  3 (bases 264 to 911)
AUTHORS    Smith,D.B., Davern,K.M., Board,P.G., Tiu,W.U., Garcia,E.G. and
           Mitchell,G.F.
TITLE      Mr 26,000 antigen of Schistosoma japonicum recognized by resistant
           WEHI 129/J mice is a parasite glutathione S-transferase [published
           erratum appears in Proc Natl Acad Sci U S A 1987 Sep;84(18):6541]
           Proc. Natl. Acad. Sci. U.S.A. 83 (22), 8703-8707 (1986)
REFERENCE  4 (bases 881 to 911)
AUTHORS    Smith,D.B., Davern,K.M., Board,P.G., Tiu,W.U., Garcia,E.G. and
           Mitchell,G.F.
TITLE      Correction: Mr 26,000 antigen of Schistosoma japonicum recognized
           by resistant WEHI 129/J mice is a parasite glutathione
           S-transferase
           Proc. Natl. Acad. Sci. U.S.A. 84, 6541-6541 (1987)
REFERENCE  5 (bases 1 to 4970)
AUTHORS    Malone,J.A.
TITLE      Direct Submission
           Submitted (19-AUG-1994) James A. Malone, International Technical
           Services, Molecular Biology Reagents Division, Pharmacia Biotech
           Inc., 2202 N. Bartlett Ave., Milwaukee, WI 53202-1009, USA
FEATURES
   source
       1..4970
       /organism="unidentified cloning vector"
       /db_xref="taxon:45196"
       /lab_host="Escherichia coli"
       183..211
       /standard_name="tac"
       /note="tac promoter for inducible expression of
       glutathione S-transferase"
       217..237
       /bound_molecy="lac repressor protein"
       258..974
       /citation=[3]
       /citation=[4]
       /codon_start=1
       /transl_table=11
       /product="glutathione S-transferase"
       /protein_id="AA057100.1"
       /db_xref="GI:595726"
       /translation="MSPIIGYWKIKGLVOPRLLEYLEKEYEELVYERDGDWRNK
       KFEIGLERPNLPYVLDGVKLTQSMALIRYIADKNNMGCGPKERAEISMEGAVIDI
       RYGVSRIVASKDFELKVDPLSKLEPMKMFEDRLCHKTYTNGVHHPDMYDALD
       VLVVMDPNCIDAFPKLVCFKKRIEAIPIQIDYIKSSKIAMPDQWATFGGCHPPK
       SDIVPRGSPGIPGSTRAAAS"
       918..935
       /note="encodes thrombin recognition site"
       930..967
       /note="Multiple Cloning Site (MCS): contains the unique
       restriction sites BamHI, EcoRI, SmaI, SalI, XhoI, and
       NotI"
       1308..2238

```

gene

```

promoter
      /gene="bla"
      /locus="1336
      /gene="bla"
      /locus="2238
      /gene="bla"
      /codon_start=1
      /transl_table=11
      /product="beta-lactamase"
      /protein_id="AA57102.1"
      /db_xref="GI:595727"
      /note="base 2996 represents the first base of the newly
      synthesized single strand"
      /direction=right
      /gene="lacIq"
      /locus="3319..4401
      /gene="lacIq"
      /codon_start=1
      /transl_table=11
      /product="lac repressor"
      /protein_id="AA57103.1"
      /db_xref="GI:595728"
      /translation="MKPYLYDAEYAGVSVOTVSRVYVNSHVSAKTRKVEAMAE
      LNYIPNVAQQLAGKQSLIGVATSLALHAPSOIVAIAKSDQLGASVSVNERS
      GVEACAAVHNILAGROVSGLIINYPIDDDQDAIAYEACNTNPALELDVSDPIINSII
      FSHSGTRIGVEHVALGHQOIALLAGLSSVABRLAGMHKLYTRNDQIARREG
      DMSAMSGFOQTOMNEGIVPTAMVLAVGADLSSABRLAGMHKLYTRNDQIARREG
      EDSSCYIPLETTIKDQFRLGOTSYDRLQLSOGAVVGNQLPLPVLRKRTTLNPT
      OTASPRALADSLMOLAROVSRLESQO"

BASE COUNT      1226 a      1202 c      1292 g      1250 t
ORIGIN
Query Match      63.6%; Score 606.6; DB 12; Length 4970;
Best local Similarity 99.2%; Pred. No. 4,6e-169;
Matches 711; Conservative 0; Mismatches 4; Indels 2; Gaps 1;

QY 1 ATGTCCCTATCTAGTATGTAATAATTAAGGCGCTTGTGAACCCACTGACTTCTT 60
    |||||||
DB 258 ATGTCCCTATCTAGTATGTAATAATTAAGGCGCTTGTGAACCCACTGACTTCTT 317
    |||||||

QY 61 TTGGAATATCTTGAAGAAAATATGAGAGCATTTTGTATGAGCGCGATGAAGTGATAA 120
    |||||||
DB 318 TTGGAATATCTTGAAGAAAATATGAGAGCATTTTGTATGAGCGCGATGAAGTGATAA 377
    |||||||

QY 121 TGGCGAAGAAAAGTTTAATTTGGGTTTGGAGTTTCCCAATCTTCTTATTTATTTGAT 180
    |||||||
DB 378 TGGCGAAGAAAAGTTTAATTTGGGTTTGGAGTTTCCCAATCTTCTTATTTATTTGAT 437
    |||||||

QY 181 GGTGATGTAATTAACACAGCTGTATGCGCATCATAGTTATATAGCTGACAGCAAC 240
    |||||||
DB 438 GGTGATGTAATTAACACAGCTGTATGCGCATCATAGTTATATAGCTGACAGCAAC 497
    |||||||

QY 241 AATGTGGGTGTTGTCCAAAGAGCGTGACAGATTTTCATGCGTTGAGAGCGGCTTTTG 300
    |||||||
DB 498 AATGTGGGTGTTGTCCAAAGAGCGTGACAGATTTTCATGCGTTGAGAGCGGCTTTTG 557
    |||||||

QY 301 GATATTAGTACGGGTGTTGCGAATGTGATATAGTAAGACTTTGAAACTCTCAAGTT 360
    |||||||
DB 558 GATATTAGTACGGGTGTTGCGAATGTGATATAGTAAGACTTTGAAACTCTCAAGTT 617
    |||||||

QY 361 GATTTTCTAGCAAGCTACGTAATGTGAAAATGTTGAGAGATGTTTATGTCATANA 420
    |||||||
DB 618 GATTTTCTAGCAAGCTACGTAATGTGAAAATGTTGAGAGATGTTTATGTCATANA 677
    |||||||

QY 421 ACATATTTAAATGGTATCATGTAAACCATCTGACTTCATGTTGATGACCGCTTTGAT 480
    |||||||

```

```

DB 678 ACATATTTAAATGGTATCATGTAAACCATCTGACTTCATGTTGATGACCGCTTTGAT 737
    |||||||
QY 481 GTTGTATTATACATGAGCAACATGTCCTGATGCGCTTCCCAAAATAGTTTGTAA 540
    |||||||
DB 738 GTTGTATTATACATGAGCAACATGTCCTGATGCGCTTCCCAAAATAGTTTGTAA 797
    |||||||

QY 541 AAACGATTTGAAGCTATCCACAAATTTGATTAAGTACTTGAATATCAGCAAGTATATAGCA 600
    |||||||
DB 798 AAACGATTTGAAGCTATCCACAAATTTGATTAAGTACTTGAATATCAGCAAGTATATAGCA 857
    |||||||

QY 601 TGGCTTTTGCAGAGCGTGGCAAGCAGCTTTGTTGGTGGCGGACCATCTTCCAAATGCGAT 660
    |||||||
DB 858 TGGCTTTTGCAGAGCGTGGCAAGCAGCTTTGTTGGTGGCGGACCATCTTCCAAATGCGAT 917
    |||||||

QY 661 CTGGTCCCGCGTGGATCCCC--GAATTCGGGCTGCACTGAGCGCGCGCTGAGA 715
    |||||||
DB 918 CTGGTCCCGCGTGGATCCCCGAGAAATTCGGGCTGCACTGAGCGCGCGCATCTGTA 974
    |||||||

RESULT 4
XXU13856
LOCUS      XXU13856      4972 bp      DNA      circular      SYN      13-DEC-1994
DEFINITION pGEX-5X1 cloning vector, complete sequence.
ACCESSION  U13856
VERSION     U13856.1 GI:595733
KEYWORDS    glutathione S-transferase; beta-lactamase; lac repressor.
SOURCE      unidentified cloning vector.
ORGANISM    unidentified sequence; vectors.
AUTHORS     1 (bases 1 to 4972)
            Malone,J.A.
TITLE       pGEX-5X1: A cloning vector for the inducible expression of genes as
            glutathione S-transferase fusion proteins containing a factor Xa
            cleavage site
            Unpublished (1994)
            2 (bases 1 to 4972)
            Smith,D.B. and Johnson,K.S.
            Single-step purification of polypeptides expressed in Escherichia
            coli as fusions with glutathione S-transferase
            Gene 67 (1), 31-40 (1988)
            88329742
            3 (bases 264 to 911)
            Smith,D.B., Davenport,K.M., Board,P.G., Tiu,W.U., Garcia,E.G. and
            Mitchell,G.F.
            Mr 26,000 antigen of Schistosoma japonicum recognized by resistant
            WEHI 129/J mice is a parasite glutathione S-transferase (published
            erratum appears in Proc Natl Acad Sci U S A 1987 Sep;84(18):6541)
            Proc. Natl. Acad. Sci. U.S.A. 83 (22), 8703-8707 (1986)
            87041520
            4 (bases 881 to 911)
            Smith,D.B., Davenport,K.M., Board,P.G., Tiu,W.U., Garcia,E.G. and
            Mitchell,G.F.
            Correction: Mr 26,000 antigen of Schistosoma japonicum recognized
            by resistant WEHI 129/J mice is a parasite glutathione
            S-transferase
            Proc. Natl. Acad. Sci. U.S.A. 84, 6541-6541 (1987)
            5 (bases 1 to 4972)
            Malone,J.A.
            Direct Submision
            Submitted (19-AUG-1994) James A. Malone, International Technical
            Services, Molecular Biology Reagents Division, Pharmacia Biotech
            Inc., 2202 N. Bartlett Ave., Milwaukee, WI 53202-1009, USA

FEATURES
    source
    promoter
    protein_bind
    1..4972
    /organism="unidentified cloning vector"
    /db_xref="taxon:45196"
    /lab_host="Escherichia coli"
    183..211
    /standard_name="lac"
    /note="lac promoter for inducible expression of
    glutathione S-transferase"
    217..237

```

```
CDS
/bound-moety-"lac repressor protein"
258..980
/citation=[3]
/citation=[4]
/codon_start=1
/transl_table=1
/product="glutathione S-transferase"
/protein_id="AA57107.1"
/db_xref="GI:595734"
/translacion="MSPILGYMKIKGVOPRLLLEYLEKEEYERHLYERDRGDKRK
KEELGEPNPIRYTIDGDKITQSMATIRIADHNMGGCPKRAETSMLEGVLDI
RYGVSRIAYSKDFETKIVDLISKLEPMKMEDELCHKRYLNGUHVHPDMLDALD
VVLVMDPCADAPFKIVCFKKRIEAPQIDKYLKSKSYIAMPLOGMOTEGGSDHPK
SDLEGRGIDPEFGRLEPRHRD"
921..932
/misc_feature
/note="encodes factor Xa recognition site"
934..969
/note="Multiple Cloning Site (MCS): contains the unique
restriction sites BamHI, EcoRI, SmaI, SalI, XhoI, and
NotI; The EcoRI site of pGEX-5x1 is in frame with the
EcoRI site of lambda gtl1"
1310..2240
/gene="bla"
/promoter
1310..1338
/gene="bla"
1380..2240
/gene="bla"
/codon_start=1
/transl_table=1
/product="beta-lactamase"
/protein_id="AA57108.1"
/db_xref="GI:595735"
/translacion="MSIOHFRVALIPFEAFLPVFAHPETLVKDAEDOLGARVY
IEDLNSGKLTIESRPERPMSTFVKVLGAVLSVADGOEOLGRIRIHSONDIYE
YSPYKEKIDGKVDKRELCATMTNTANILITIGKEKELTALRHNGCHVRL
DRWPELNEALPDEROTTPMAAMATLRKLITELTLASRQQLIDMEEDRKAGPL
LRSLPAGWFIADKSGAGERSGRIIALGDPKPSRIVITVTSQATMDERNQILA
EIGASLIRKM"
2998
/rep_origin
/note="base 2998 represents the first base of the newly
synthesized single strand"
/direction=right
3321..4403
/gene="lacIq"
3321..4403
/gene="lacIq"
/codon_start=1
/transl_table=1
/product="lac repressor"
/protein_id="AA57109.1"
/db_xref="GI:595736"
/translacion="MKPVTLIDVAEYAGVSQTVSRVYNOASHVSAKTRKEVAAAE
LNTIPNVAQOLAKQKSLITGVATSSIALINPQIVAAIKRAQLGASVYNSVENS
GVEACKAVNHLAQRVSGILINPLDDODAIAEACTNVPALPLVSDQTPINSII
FSHEDGRLAEVHLVALGHOQIALAGPLSSVSARLRIAGMKYLTNRQIPIAREG
DWSMSGFOQDMQINIEGIYPTAMLVANDQMALGAMRAITSGRVAADISVYDPT
EDSSCYIPPLTIKODPRILGOTSVDRILQSGQAVKGNQLLPVSLVKKRTTLPNT
QTASPRALASIMOLAROVSLTESGO"
BASE COUNT 1228 a 1201 c 1293 g 1250 t
ORIGIN
```

```
Query Match 62.5% Score 685; DB 12; Length 4972;
Best Local Similarity 98.1%; Pred. No. 4.6e-166;
Matches 705; Conservative 0; Mismatches 10; Indels 4; Gaps 1;
```

```
QY 1 ATGCCCCCTACTAGTATTTGAAGGAGCCCTTGCAACCACTGACTCTT 60
|||||
DB 258 ATGCCCCCTACTAGTATTTGAAGGAGCCCTTGCAACCACTGACTCTT 317
|||||
QY 61 TTGAATATCTTGAGAAATAATGAAGCATTTGTATGAGCGGAGTGAAGTATAA 120
|||||
DB 318 TTGAATATCTTGAGAAATAATGAAGCATTTGTATGAGCGGAGTGAAGTATAA 377
|||||
```

```
QY 121 TGGCAACAAAGAAAGTTGAATGGGTTTGAGATTGCCAATCTCTTATATATGAT 180
|||||
DB 378 TGGCAACAAAGAAAGTTGAATGGGTTTGAGATTGCCAATCTCTTATATATGAT 437
|||||
QY 181 GGTGATGTTAAATTAACACAGCTATAGCCCATCATAGCTTATATAGTGACACACAA 240
|||||
DB 438 GGTGATGTTAAATTAACACAGCTATAGCCCATCATAGCTTATATAGTGACACACAA 497
|||||
QY 241 ATGTTGGGTGGTGTGCCAAAGAGCGTGACAGATTTAAAGCTTGAAGAGCGGTTTG 300
|||||
DB 498 ATGTTGGGTGGTGTGCCAAAGAGCGTGACAGATTTAAAGCTTGAAGAGCGGTTTG 557
|||||
QY 301 GATATTAGATACAGCTGTTTCCAGAAATTCATATAGTAAGAATTGAAACTCTCAAGTT 360
|||||
DB 558 GATATTAGATACAGCTGTTTCCAGAAATTCATATAGTAAGAATTGAAACTCTCAAGTT 617
|||||
QY 361 GATTTTCTTAGCAACGTAACGTAAGTGTGAAGATGTTGAGATGTCATATAA 420
|||||
DB 618 GATTTTCTTAGCAACGTAACGTAAGTGTGAAGATGTTGAGATGTCATATAA 677
|||||
QY 421 ACATATTAAATGATGATCATATAGTATACCATGCTGACTTCATAGTTATGACGCTTGAT 480
|||||
DB 678 ACATATTAAATGATGATCATATAGTATACCATGCTGACTTCATAGTTATGACGCTTGAT 737
|||||
QY 481 GTTGTATTATACATGACCAACCAATGCTGATGCTGCCAAATATGTTGTTTAAA 540
|||||
DB 738 GTTGTATTATACATGACCAACCAATGCTGATGCTGCCAAATATGTTGTTTAAA 797
|||||
QY 541 AAACGTAATGAAGCTATATCCACAAATTTAATAAGTATGAAATCCAGCAAGTATATACA 600
|||||
DB 798 AAACGTAATGAAGCTATATCCACAAATTTAATAAGTATGAAATCCAGCAAGTATATACA 857
|||||
QY 601 TGGCCTTGCGAGGCGTGCGAAGCCAGCTTGGTGGTGCGACCATCTCCAAATTCGGAT 660
|||||
DB 858 TGGCCTTGCGAGGCGTGCGAAGCCAGCTTGGTGGTGCGACCATCTCCAAATTCGGAT 917
|||||
QY 661 CTGTTCCGCGT---GGATCCCGCAATTCGCGGCTGACATCGAGCGCGCGCTGAGA 715
|||||
DB 918 CTGATCGAAGGTCGTGGATCCCGCAATTCGCGGCTGACATCGAGCGCGCGCTGAGA 976
|||||
```

```
RESULT 5
XXU13849
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
REFERENCE
```

XXU13849 4947 bp DNA circular SYN 13-DEC-1994
pGEX-1 Lambda T cloning vector, complete sequence.
U13849
U13849.1 GI:595705
glutathione S-transferase; beta-lactamase; lac repressor.
unidentified cloning vector.
unidentified cloning vector.
artificial sequence; vectors.
1 (bases 1 to 4947)
Malone,J.A.
pGEX-1 Lambda T: A cloning vector for the inducible expression of
genes in glutathione S-transferase fusion proteins containing a
thrombin cleavage site
Unpublished (1994)
2 (bases 1 to 4947)
Smith,D.B., and Johnson,K.S.
Single-step purification of polypeptides expressed in Escherichia
coli as fusions with glutathione S-transferase
Gene 67 (1), 31-40 (1988)
88329742
3 (bases 264 to 911)
Smith,D.B., Davenport,K.M., Board,P.G., Tlu,W.U., Garcia,E.G. and
Mitchell,G.F.
Mr 26,000 antigen of Schistosoma japonicum recognized by resistant
WEHI 129/7 mice is a parasite glutathione S-transferase [published
erratum appears in Proc Natl Acad Sci U S A 1987 Sep;84(18):6541]
JOURNAL
MEDLINE
PROC. NATL. ACAD. SCI. U.S.A. 83 (22), 8703-8707 (1986)
87041520
4 (bases 881 to 911)

Query Match	Best Local Similarity	Matches	Score	Pred.	Mismatches	Indels	Gaps	Length
62.3%	99.1%	0	682.4	2.1e-165	6	0	0	4947
Conservative	0	0	6	0	0	0	0	0
1	ATGTCCTCTATACAGAGTTATTTGGAAATTAAGAGGCTTTGCAACCCATCGACTTCTT	60						
258	ATGTCCCTTACTAGATTTATTTGGAATTAAGGCTTTGCAACCCATCGACTTCTT	317						
61	TTGCAATATCTTGAAGAAAAATATGACAGCATTTGTATGAGCCGATGAAGTGATAAA	120						
318	TTGCAATATCTTGAAGAAAAATATGACAGCATTTGTATGAGCCGATGAAGTGATAAA	377						
121	TGGCGAACAACAAAGTTGATTTGGTGGTGGATTTCCCATCTTCTTATATATGAT	180						
378	TGGCGAACAACAAAGTTGATTTGGTGGTGGATTTCCCATCTTCTTATATATGAT	437						
181	GGTATGTTAAATTAACACAGCTCTATGGCCATCATACGTTATATAGCTGACACACAC	240						
438	GGTATGTTAAATTAACACAGCTCTATGGCCATCATACGTTATATAGCTGACACACAC	497						
241	ATGTTGGGTGGTGTGCCAAAAGCGGTGCAGAGATTTCAATGCTTTGAAGAGCGGTTTG	300						
498	ATGTTGGGTGGTGTGCCAAAAGCGGTGCAGAGATTTCAATGCTTTGAAGAGCGGTTTG	557						
301	GATATTAATACGCGTGTTCGAGAAATTCGATTAAGCTTTGAAGCTTCGAAAGTT	360						
558	GATATTAATACGCGTGTTCGAGAAATTCGATTAAGCTTTGAAGCTTCGAAAGTT	617						
361	GATTTCTTACAGAGCTACCTGAAATGCTGAAAGATGCTGAAAGATGCTTATGTCATAAA	420						
618	GATTTCTTACAGAGCTACCTGAAATGCTGAAAGATGCTGAAAGATGCTTATGTCATAAA	677						
421	ACATATTTAAATGCGATCATGTAAACCCATCCTGACTTCATGTTGTATGAGCTCTTGAT	480						
678	ACATATTTAAATGCGATCATGTAAACCCATCCTGACTTCATGTTGTATGAGCTCTTGAT	737						
481	GTTGTTTATATACATGACCCCAATGTGCTGCGATGCGTTCCCAAAATTAAGTTGTTTAAA	540						
738	GTTGTTTATATACATGACCCCAATGTGCTGCGATGCGTTCCCAAAATTAAGTTGTTTAAA	797						
541	AAACGATTTGAAGCTATCCCAACAAATTTGATTAAGTACTGGAANTCCAGACAGATATATACA	600						
798	AAACGATTTGAAGCTATCCCAACAAATTTGATTAAGTACTGGAANTCCAGACAGATATATACA	857						
601	TGGCCCTTTCGAGGCGTGCAGAACCCAGCTTTGGTGTGCGACCATCTCCCAAAATCGGAT	660						
858	TGGCCCTTTCGAGGCGTGCAGAACCCAGCTTTGGTGTGCGACCATCTCCCAAAATCGGAT	917						
661	CTGGTTCGCGGTGGATCCCGCATTCGCCGGGT	692						
918	CTGGTTCGCGGTGGATCCCGCATTCATCGT	949						

BASE COUNT: 1223 a 1193 c 1283 g 1248 t

ORIGIN

/codon_start=1

/transl_table=1

/product="Iac repressor"

/protein_id="AA57088.1"

/db_xref="GI:595708"

/translation="MKPVTLLYDVAEYAGVSYOTVSRVYNASHSAKTRPKVFAAAE
LNTIPNRVADQQLAGKOSLLIGVATSSILALHPQSYVAATIKSRDQJGASVYVSWERS
GVCEKAAVHNLLAORVSGLIINPLDDQDAIVEACTVNPALFLDVSDQTPINSLI
FSHEDGRLEVEHLVALGHQOIALVLNGPSSVSAKRLACMHKLTFRNQIOPIAERGG
DMSAGSFQOTKQMLNEGIVPTAMLVANPDMAIGAMALATIESLRVADISVYGDF
EDSSCYPLPTTIIKODRRLGOTSVDBLLQDSQAVKQNDLLPVSIVKRTTLLAPNT
OTFSPRLAASIMQIALAQVSRLESCQ"

RESULT 6

AB014641

LOCUS

DEFINITION

3417 bp DNA circular SYN 30-SEP-1998

Cloning vector pGEX-PUC-3P DNA, complete sequence.

ACCESSION AB014641.1 GI:3184403
 VERSION AB014641.1 GI:3184403
 KEYWORDS Cloning vector: GST-stuffer protein; GST-stuffer fusion protein.
 SOURCE Cloning vector pGEX-PUC-3T DNA.
 ORGANISM Cloning vector pGEX-PUC-3T
 Cloning vector pGEX-PUC-3T
 artificial sequence: vectors.
 REFERENCE 1 (bases 1 to 3417)
 AUTHORS Fukunaga, R.
 TITLE Submitted (22-MAY-1998) to the DDBJ/EMBL/GenBank databases. Rikio
 JOURNAL Osaka University Medical School, Department of Genetics
 Fukunaga, Osaka University Medical School, Department of Genetics
 B-3; 2-2 Yamadaoka, Suita, Osaka 565-0871, Japan
 (E-mail: fukunaga@genetic.med.osaka-u.ac.jp, Tel: 81-6-879-3318,
 Fax: 81-6-879-3319)
 2 (sites)
 REFERENCE Fukunaga, R. and Hunter, T.
 AUTHORS MNK1, a new MAP kinase-activated protein kinase, isolated by a
 TITLE novel expression screening method for identifying protein kinase
 JOURNAL substrates
 MEDLINE EMBO J. 16 (8), 1921-1933 (1997)
 FEATURES 97299869
 SOURCE Location/Qualifiers
 1. .3417
 /organism="Cloning vector pGEX-PUC-3T"
 /db_xref="taxon:77040"
 9. .1997
 /note="pUC119"
 misc_feature 2005. .2953
 /note="pGEX-KG"
 misc_feature 2246. .3142
 /gene="GST-Stuffer fusion"
 2246. .3142
 /gene="GST-Stuffer fusion"
 /codon_start=1
 /transl_table=1
 /product="GST-stuffer fusion protein"
 /protein_id="BAA28713.1"
 /db_xref="GI:3184404"
 /translation="MSPILGVYKIKGVOPRILLEYLEKEEHLERDEGDKRMK
 KFEUGLEFPMPLPYIIDGVDLTQSMALIRIADKHNMLGCKPRRAISMIEGVLDI
 RYGVSRATSKDEPTLKVDFLSKLEPMKEFDLCKITLUNGSHVHPDMLDALD
 VLKMDPMDLADPEKRLVCFKRIEAIPOIDIKYKSSVIAMPLQGWATGCGHPK
 SDIYRSGSPGISGGGCGICPGAGLEGILLHDCGHPGAQAAVSWWDTEGILPYGAAT
 EPGSTRAGEQOEAEQAOODV"
 2968. .3385
 /note="Stuffer: that the stuffer DNA is derived from the
 420bp PvuII fragment(1802-2221) of the human G-CSF
 receptor cDNA (Acc# No. M38025)"
 misc_feature 2968. .3385
 /note="Stuffer: that the stuffer DNA is derived from the
 420bp PvuII fragment(1802-2221) of the human G-CSF
 receptor cDNA (Acc# No. M38025)"
 BASE COUNT 864 a 781 c 898 g 874 t
 ORIGIN
 Query Match 62.1%; Score 681; DB 12; Length 3417;
 Best Local Similarity 100.0%; Pred. No. 4,7e-165;
 Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 ATGCCCTACTAGTGGTTATGGAATTAAGGCGCTTGCAACCCACGACTTCT 60
 |||||||
 Db 2246 ATGCCCTACTAGTGGTTATGGAATTAAGGCGCTTGCAACCCACGACTTCT 2305
 |||||||
 Oy 61 TTGCAATCTTGAAGAAATATGAGAGCATTTGTATGAGCCGATGAAGTGATAAA 120
 |||||||
 Db 2306 TTGCAATCTTGAAGAAATATGAGAGCATTTGTATGAGCCGATGAAGTGATAAA 2365
 |||||||
 Oy 121 TGGCGAACAACAAGTTGAATGGGTTTGAAGTTCCCAATCTTCTATTATATTGAT 180
 |||||||
 Db 2366 TGGCGAACAACAAGTTGAATGGGTTTGAAGTTCCCAATCTTCTATTATATTGAT 2425
 |||||||
 Oy 181 GGTATGTTAATTAACACATCTATGGCATCTATACGTTATATAGTCGACAGACAC 240
 |||||||
 Db 2426 GGTATGTTAATTAACACATCTATGGCATCTATACGTTATATAGTCGACAGACAC 2485
 |||||||
 Oy 241 ATGTGGGTGGTGTCCAAAAGACGCTGACAGATTTCATGCTTGAGAGACGCGTTTG 300

|||||
 Db 2486 ATGTGGGTGGTGTCCAAAAGACGCTGACAGATTTCATGCTTGAGAGACGCGTTTG 2545
 |||||||
 Oy 301 CATATTAAATACGGGTGTTTCGAGAAATTCGATATAGTAAGACTTTGAAACTCTAAAGTT 360
 |||||||
 Db 2546 GATATTAGATACGGGTGTTTCGAGAAATTCGATATAGTAAGACTTTGAAACTCTAAAGTT 2605
 |||||||
 Oy 361 GATTTCCTTACCAAGCTACCTGAAGTGTGAAGAAATGTTGGAAGATCGTTTATGATATAA 420
 |||||||
 Db 2606 GATTTCCTTACCAAGCTACCTGAAGTGTGAAGAAATGTTGGAAGATCGTTTATGATATAA 2665
 |||||||
 Oy 421 ACAATATTAATGATGATCATGTAAACCATCCATCCGATCTCATCTGTATAGACGCTTTGAT 480
 |||||||
 Db 2666 ACAATATTAATGATGATCATGTAAACCATCCATCCGATCTCATCTGTATAGACGCTTTGAT 2725
 |||||||
 Oy 481 GTTGTTTATACATGAGACCAATGCTGCTGATGATGTTGCCCAAAATGTTGTTTAAA 540
 |||||||
 Db 2726 GTTGTTTATACATGAGACCAATGCTGCTGATGATGTTGCCCAAAATGTTGTTTAAA 2785
 |||||||
 Oy 541 AAAGTATTAAGCTATCCCAAAATGATAGTAACTTGAATCCAGACAGATATATAGCA 600
 |||||||
 Db 2786 AAAGTATTAAGCTATCCCAAAATGATAGTAACTTGAATCCAGACAGATATATAGCA 2845
 |||||||
 Oy 601 TGGCCTTTGCAAGGCTGCGCAAGCCAGCTTTGTGTGTGCGGACCATCTCCAAATCGGAT 660
 |||||||
 Db 2846 TGGCCTTTGCAAGGCTGCGCAAGCCAGCTTTGTGTGTGCGGACCATCTCCAAATCGGAT 2905
 |||||||
 Oy 661 CTGCTCCGCGCTGATCCCG 681
 |||||||
 Db 2906 CTGCTCCGCGCTGATCCCG 2926
 |||||||

RESULT 7
 LOCUS XXU13850 4948 bp DNA circular SYN 13-DEC-1994
 DEFINITION pGEX-2T cloning vector, complete sequence.
 ACCESSION U13850
 VERSION U13850.1 GI:595709
 KEYWORDS glutathione S-transferase; beta-lactamase; lac repressor.
 SOURCE unidentified cloning vector.
 ORGANISM unidentified cloning vector.
 REFERENCE 1 (bases 1 to 4948)
 AUTHORS Malone, J. A.
 TITLE pGEX-2T: A cloning vector for the inducible expression of genes as
 glutathione S-transferase fusion proteins containing a thrombin
 cleavage site
 JOURNAL Unpublished (1994)
 REFERENCE 2 (bases 1 to 4948)
 AUTHORS Smith, D.B. and Johnson, K.S.
 TITLE Single-step purification of polypeptides expressed in *Escherichia*
 coli as fusions with glutathione S-transferase
 JOURNAL Gene 67 (1), 31-40 (1988)
 MEDLINE 88329742
 REFERENCE 3 (bases 264 to 911)
 AUTHORS Smith, D.B., Davern, K.M., Board, P.G., Tlu, W.U., Garcia, E.G. and
 Mitchell, G.F.
 TITLE Mr 26,000 antigen of *Schistosoma japonicum* recognized by resistant
 WEHI 129/J mice is a parasite glutathione S-transferase [published
 erratum appears in Proc Natl Acad Sci U S A 1987 Sep;84(18):6541]
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 83 (22), 8703-8707 (1986)
 MEDLINE 87041520
 REFERENCE 4 (bases 881 to 911)
 AUTHORS Smith, D.B., Davern, K.M., Board, P.G., Tlu, W.U., Garcia, E.G. and
 Mitchell, G.F.
 TITLE Correction: Mr 26,000 antigen of *Schistosoma japonicum* recognized
 by resistant WEHI 129/J mice is a parasite glutathione
 S-transferase
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 84, 6541-6541 (1987)
 REFERENCE 5 (bases 1 to 4948)
 AUTHORS Malone, J. A.
 TITLE Direct Submission
 JOURNAL Submitted (19-AUG-1994) James A. Malone, International Technical

Services, Molecular Biology Reagents Division, Pharmacia Biotech
Inc., 2202 N. Bartlett Ave., Milwaukee, WI 53202-1009, USA

FEATURES

source

1. 4948
Location/Qualifiers

/organism="unidentified cloning vector"
/db_xref="taxon:45196"

/lab_host="Escherichia coli"

promoter

183. 211
/standard_name="tac"
/note="tac promoter for inducible expression of
glutathione S-transferase"

protein_bind

217. 237
/gene="lac"
/bound_molecule="lac repressor protein"

gene

217. 237
/gene="lac"

CDS

258. 956
/citation={3}
/citation={4}
/codon_start=1
/transl_table=11
/product="glutathione S-transferase"

/protein_id="AA57089.1"
/db_xref="GI:595710"
/translation="MSPILGYWKIKGLVQPRLLLEEKYEHLRYERDEGKMRNK
KFELEPPNLPYIDGDVFKLOSMAIRIYIDKINMKGGCKEAEHISMEGAVIDI
RYGVRIRAYSKDEFTKLVDFKLSKPEMKHCHYUHGHDHIDPDHLYALD
VLYMDPNCIDAFPLVCFKKRIEAIPIQIDTKSKSYIAMPLQGMQATFFGGDHPK
SDIYRGSPIHRD"

misc_feature

misc_feature

918. 935
/note="encodes thrombin recognition site"

misc_feature

930. 945
/note="Multiple Cloning Site (MCS): contains the unique
restriction sites BamHI, SmaI and EcoRI"

gene

promoter

CDS

1286. 2216
/gene="bla"
1286. 1314
/gene="bla"
1356. 2216
/gene="bla"
/codon_start=1
/transl_table=11
/product="beta-lactamase"
/protein_id="AA57090.1"
/db_xref="GI:595711"
/translation="MSIOHFRVALIPFAAPCLPYFAHPENTLVKDAEDOLGARVY
IETDLSKILIESFRPERFRPMSTFKVLLCGAVSRVDAGGEOLGRRIHYSQNDLVE
YSPVTEKLTIDMTVRELCSAIIWSDMTAANLLITGGPRELTAFLHNMGDHVTPL
DRWEPELNEALPNDERTTMAPAAMATTLRLKLTGLTLLASRQQLIDWMEADKVGPL
LRSAIPACMFADKSGAGERSGIIAALGPDGKPSRIIVYITGSOATMDERNQIA
EIGASLIKHM"

rep_origin

2974
/note="base 2974 represents the first base of the newly
synthesized single strand"

gene

CDS

3297. 4379
/direction=right
/gene="lacIq"
3297. 4379
/gene="lacIq"
/codon_start=1
/transl_table=11
/product="lac repressor"
/protein_id="AA57091.1"
/db_xref="GI:595712"
/translation="MKPYTLVDVAYAGSYQTVSRVYNQASHVSATREKYEAMAE
LNTIPNRVAQOLAGKOSLLIGVATSSLALHAPSQVAATKSRADOLGASVSVSMERS
EVEACKAIVNHLAQRVSGLIINPLDQDAVAACQTNVPLDLVSDQPIINSII
FSHEDGTRGLVGHVLAHQOIALLAGPDSVSAVRLAGMHRKYLTRNOIQIARREG
DMSASGFOOTQOMNEGIVPTAMLVANDOMLGAAMRAITGESVAGADISVGDYDT
EDSSCYIPDLTIKODPRLGOTSVDRLLQLLSQGAIVGNQLLPVSLVKRRKTTLPNT
QTASPRALADSLMOAROVSRLESQ"

BASE COUNT

1223 a 1193 c 1284 g 1248 t

Query Match 62.1%; Score 681; DB 12; Length 4948;
Best Local Similarity 100.0%; Pred. No. 4,9e-165;
Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTCCCTATACCTAGGATTTGGAAAATTTAGGGCTTGTGCACCCACTGCATCTT 60
|||||
Db 258 ATGTCCCTATACCTAGGATTTGGAAAATTTAGGGCTTGTGCACCCACTGCATCTT 317
Qy 61 TTGGATATCTTGAAGAAAAATATGAAGACATTTGTATAGCGCATGAGGTGATAA 120
|||||
Db 318 TTGGAATATCTTGAAGAAAAATATGAAGACATTTGTATAGCGCATGAGGTGATAA 377
Qy 121 TGGCGAACAACAAAGCTTTCGATGGGTTTGGAGTTTCCCAATCTTCTATTTATGAT 180
|||||
Db 378 TGGCGAACAACAAAGCTTTCGATGGGTTTGGAGTTTCCCAATCTTCTATTTATGAT 437
Qy 181 GGTGATGTTAAATTAACACAGTCTATGCGCATCATGATATATAGCTGACAGCACAC 240
|||||
Db 438 GGTGATGTTAAATTAACACAGTCTATGCGCATCATGATATATAGCTGACAGCACAC 497
Qy 241 ATGTGGGTTGGTTTCCCAAAAGCGCTGACAGATTTCAATGCTTGAAGAGCGGTTTG 300
|||||
Db 498 ATGTGGGTTGGTTTCCCAAAAGCGCTGACAGATTTCAATGCTTGAAGAGCGGTTTG 557
Qy 301 GATATTAGATACGCTGTTTCGAGAAATTCATATAGTAAGACTTTGAACCTGCAAGTT 360
|||||
Db 558 GATATTAGATACGCTGTTTCGAGAAATTCATATAGTAAGACTTTGAACCTGCAAGTT 617
Qy 361 GATTTTCTTAGCAAGCTACCTGAATGCTGAAAATGTTGCAAGATGTTATGTCATAA 420
|||||
Db 618 GATTTTCTTAGCAAGCTACCTGAATGCTGAAAATGTTGCAAGATGTTATGTCATAA 677
Qy 421 ACATATTTAAATGCTATCATATACCATCTGACTTCAATGTTATGACGCTTTGAT 480
|||||
Db 678 ACATATTTAAATGCTATCATATACCATCTGACTTCAATGTTATGACGCTTTGAT 737
Qy 481 GTTGTTTATACATGACCAACATGCTGCGATGCGTCCCAAAATAGTTGTTTAA 540
|||||
Db 738 GTTGTTTATACATGACCAACATGCTGCGATGCGTCCCAAAATAGTTGTTTAA 797
Qy 541 AAACGATTTGAACCTATCCCAAAATGATAGTACTTGAATCCAGCAAGTATATAGCA 600
|||||
Db 798 AAACGATTTGAACCTATCCCAAAATGATAGTACTTGAATCCAGCAAGTATATAGCA 857
Qy 601 TGGCCTTTGACAGGCTGCGACAGCCAGCTTTGTTGTTGCGGACCATCTCCAAAATGGAT 660
|||||
Db 858 TGGCCTTTGACAGGCTGCGACAGCCAGCTTTGTTGTTGCGGACCATCTCCAAAATGGAT 917
Qy 661 CTGCTCCGCGTGATCCCG 681
|||||
Db 918 CTGCTCCGCGTGATCCCG 938

RESULT 8

EV084572

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

TITLE

AUTHORS

REFERENCE

TITLE

EV084572 5042 bp DNA circular SYN 04-FEB-1997
Expression vector pGEX-2T-6His-PL2 (GST) gene, complete
cds, complete vector sequence.

084572.1 GI:1814367

Expression vector pGEX-2T-6His-PL2.
Expression vector pGEX-2T-6His-PL2.
artificial sequence; vectors.

1 (bases 1 to 5042)
Hipskind, R. H. and Delaney, P.

Unpublished
journal

2 (bases 1 to 5042)
Cahill, M. A.

Direct Submission

JOURNAL

Submitted (09-JAN-1997) Immunology & Cell Biology, JCSMR, Australian National University, Acton, ACT 2601, Australia
 Constructed by Patricia Delaney and Bob Hipskind, Institute for Molecular Biology, Hannover Medical School, 30623 Hannover, Germany, June 1991. Parent = pGEX-2T (Pharmacia). Submitted by Mike Cahill.

FEATURES

source

Location/Qualifiers
 1..5042
 /organism="Expression vector pGEX-2T-6His-PL2"
 /db_xref="taxon:56446"
 /note="pGEX-2T derivative; pGEX plasmid"
 258..1046
 /gene="GST"
 /note="GST-6his-PL2"
 /allele="GST-2T-6his-PL2"
 258..1046
 /gene="GST"
 /codon_start=1
 /product="GST-6his"
 /protein_id="AAB41883.1"
 /db_xref="GI:1814368"
 /translation="MSPILGYWKIKGLVOPRLLLEYLEEKEEHLVERDEGRKRNK
 KEFLGLEPNNPYITDGDVKLTOSMAITRYIAKHNNLGGCPKRAEISMLEGAVLDI
 RYGVSRVAYSKEDEFLKVDPLSKLPEMKNEEDRLCHKTYLNGDHTVHPDMLDALD
 VVLYMDPMDLAFPKLVCFKRIEAIPOIDKYLKSSKYIAMPLOGMOATPFGGDPHPK
 SDLVPRGSPVEMSHHHHNGEFRESSISHAVPLEAMDPTNSS"

CDS

BASE COUNT 1245 a 1223 c 1309 g 1265 t
 ORIGIN

Query Match 62.1%; Score 681; DB 12; Length 5042;
 Best Local Similarity 100.0%; Pred. No. 4.9e-165;
 Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGCCCCCTACTAGTATTTGGAATTTAAGGCGCTTGACACCCAGCTGACTCT 60
 258 ATGTCCTCTATCTAGTATTTGGAATTTAAGGCGCTTGACACCCAGCTGACTCT 317
 61 TTGAATATCTTGAAGAAAATATGAAGAGCATTTGTATGAGCGGATGAAGTATAA 120
 318 TTGAATATCTTGAAGAAAATATGAAGAGCATTTGTATGAGCGGATGAAGTATAA 377
 121 TGGCAAAACAAAATTTGAATTTGGGTTTCCCAATCTTCCTATTATATTTGAT 180
 378 TGGCAAAACAAAATTTGAATTTGGGTTTCCCAATCTTCCTATTATATTTGAT 437
 181 GGTGATTTAATTAACACAGCTATGCGCATCATAGCTTATATAGTACAGACACAC 240
 438 GGTGATTTAATTAACACAGCTATGCGCATCATAGCTTATATAGTACAGACACAC 497
 241 ATGTTGGGTGTTGCAAAAAGAGGTGACAGATTTCAATGCTTGAAGAGCGGTTTG 300
 498 ATGTTGGGTGTTGCAAAAAGAGGTGACAGATTTCAATGCTTGAAGAGCGGTTTG 557
 301 GATATTAATACGCTGTTTCAGAAATTCATATAGTAAAGACTTTGAACCTCTCAAGTT 360
 558 GATATTAATACGCTGTTTCAGAAATTCATATAGTAAAGACTTTGAACCTCTCAAGTT 617
 361 GATTTTCTTACCAAGCTACCTGAATGCTGAAGATGTCGAAGATCGTTTATGATATAA 420
 618 GATTTTCTTACCAAGCTACCTGAATGCTGAAGATGTCGAAGATCGTTTATGATATAA 677
 421 ACATATTTAATGATGATGATGAACCATCTGACTTCATGTTGATAGAGCTCTTGAT 480
 678 ACATATTTAATGATGATGATGAACCATCTGACTTCATGTTGATAGAGCTCTTGAT 737
 481 GTTGTATATCATGAGCAACCATGCTGATGCTGCTTCCCAAAATTAAGTTGTTTTAA 540
 738 GTTGTATATCATGAGCAACCATGCTGATGCTGCTTCCCAAAATTAAGTTGTTTTAA 797
 541 AAAGCATTTGAAGCTATCCCAAAATTTGATAGTAAAGTGAAGTACAGAAATATATGCA 600
 798 AAAGCATTTGAAGCTATCCCAAAATTTGATAGTAAAGTGAAGTACAGAAATATATGCA 857

QY 601 TGGCCTTTGACAGGCGTGGCAAGCCAGCTTTGGTGGCGACCATCTCCAAATCGGAT 660
 DB 858 TGGCCTTTGACAGGCGTGGCAAGCCAGCTTTGGTGGCGACCATCTCCAAATCGGAT 917
 QY 661 CTGGTTCCGCGTGAATCCCG 681
 DB 918 CTGGTTCCGCGTGAATCCCG 938

RESULT 9

EV084571

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

1 (bases 1 to 5052)
 Expression vector pGEX-2T-6H.
 Expression vector pGEX-2T-6H.
 artificial sequence: vectors.
 1 (bases 1 to 5052)
 Direct Submission
 Submitted (09-JAN-1997) Immunology & Cell Biology, JCSMR, Australian National University, Acton, ACT 2601, Australia
 Constructed by Mike Cahill, Institute for Molecular Biology, Hannover Medical School, 30623 Hannover, Germany, April 1995.
 Parent pGEX-2T-6his-PL2, (pGEX-2T derivative). Additional Bgl II, BsrG I, Nde I and Xba I sites to polylinker. Also called pGEX-2T-6H-BX#2.

FEATURES

source

Location/Qualifiers
 1..5052
 /organism="Expression vector pGEX-2T-6H"
 /db_xref="taxon:56445"
 /note="pGEX-2T derivative; pGEX plasmid"
 258..1064
 /gene="GST"
 /note="GST-6his"
 /allele="GST-2T-6his"
 258..1064
 /gene="GST"
 /codon_start=1
 /product="GST-6his"
 /protein_id="AAB41882.1"
 /db_xref="GI:1814368"
 /translation="MSPILGYWKIKGLVOPRLLLEYLEEKEEHLVERDEGRKRNK
 KEFLGLEPNNPYITDGDVKLTOSMAITRYIAKHNNLGGCPKRAEISMLEGAVLDI
 RYGVSRVAYSKEDEFLKVDPLSKLPEMKNEEDRLCHKTYLNGDHTVHPDMLDALD
 VVLYMDPMDLAFPKLVCFKRIEAIPOIDKYLKSSKYIAMPLOGMOATPFGGDPHPK
 SDLVPRGSPVEMSHHHHNGEFREICITYELSRACGTSRGSLSGVDPRIYD"

CDS

BASE COUNT 1250 a 1224 c 1308 g 1270 t
 ORIGIN

Query Match 62.1%; Score 681; DB 12; Length 5052;
 Best Local Similarity 100.0%; Pred. No. 4.9e-165;
 Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGCCCCCTACTAGTATTTGGAATTTAAGGCGCTTGACACCCAGCTGACTCT 60
 258 ATGTCCTCTATCTAGTATTTGGAATTTAAGGCGCTTGACACCCAGCTGACTCT 317
 61 TTGAATATCTTGAAGAAAATATGAAGAGCATTTGTATGAGCGGATGAAGTATAA 120
 318 TTGAATATCTTGAAGAAAATATGAAGAGCATTTGTATGAGCGGATGAAGTATAA 377
 121 TGGCAAAACAAAATTTGAATTTGGGTTTCCCAATCTTCCTATTATATTTGAT 180
 378 TGGCAAAACAAAATTTGAATTTGGGTTTCCCAATCTTCCTATTATATTTGAT 437
 181 GGTGATTTAATTAACACAGCTATGCGCATCATAGCTTATATAGTACAGACACAC 240

Db 438 GGTGATGTTAAATTAACACAGCTTATGCCATCATACGTTATATAGCTGACACACAC 497
 Qy 241 ATGTTGGGTGTTGTCACAAAGAGCGTGCAGAGATTTCAATGCTTGAAGAGCGGTTTG 300
 Db 498 ATGTTGGGTGTTGTCACAAAGAGCGTGCAGAGATTTCAATGCTTGAAGAGCGGTTTG 557
 Qy 301 GATATTAGTACGCTGTTTCGAGAAATTCATATAGTAAGACTTTGAACCTCTCAAGTT 360
 Db 558 GATATTAGTACGCTGTTTCGAGAAATTCATATAGTAAGACTTTGAACCTCTCAAGTT 617
 Qy 361 GATTTCTTAGACAGCTACCTGAATGCGAATGTTGGAAGATGTTATGCTCAATAA 420
 Db 618 GATTTCTTAGACAGCTACCTGAATGCGAATGTTGGAAGATGTTATGCTCAATAA 677
 Qy 421 ACATATTTAAATGATGATCATGTAACCCATCCGACTTCATGTTGATGACGCTTGAT 480
 Db 678 ACATATTTAAATGATGATCATGTAACCCATCCGACTTCATGTTGATGACGCTTGAT 737
 Qy 481 GTTGTGTTTATACATGACCAATGTCCTGATGCGTTCACCAAAATTAAGTTGTTTAA 540
 Db 738 GTTGTGTTTATACATGACCAATGTCCTGATGCGTTCACCAAAATTAAGTTGTTTAA 797
 Qy 541 AAACGATTTAGAGCTATGCCACAAATGATAGTACTTGAATCCAGCAAGTATATAGA 600
 Db 798 AAACGATTTAGAGCTATGCCACAAATGATAGTACTTGAATCCAGCAAGTATATAGA 857
 Qy 601 TGCGCTTTGACAGGCGTGGCAGGCAAGCTTTGATGATGCGGACCATCTCCAAATGCGAT 660
 Db 858 TGCGCTTTGACAGGCGTGGCAGGCAAGCTTTGATGATGCGGACCATCTCCAAATGCGAT 917
 Qy 661 CTGCTTCGCGGTGATCCCG 681
 Db 918 CTGCTTCGCGGTGATCCCG 938

RESULT 10

LOCUS EV085201 5501 bp DNA circular SYN 31-MAR-1998
 DEFINITION Expression vector pGH/F2.2-163, complete sequence.
 ACCESSION U85201
 VERSION U85201.1 GI:3002495

KEYWORDS Expression vector pGH/F2.2-163.
 ORGANISM Expression vector pGH/F2.2-163.
 SOURCE Expression vector pGH/F2.2-163.
 REFERENCE 1 (bases 1 to 5501)
 AUTHORS Cahill, M.A.
 TITLE Phosphorylation of Transcription Factors Elk-1 and Fra-2 by Associated Serum-dependent Kinase Activity
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 5501)
 AUTHORS Cahill, M.A.
 TITLE Direct Submission
 JOURNAL Submitted (14-JAN-1997) Division of Immunology & Cell Biology, JCSMR, Australia National University, P.O. Box 334, Canberra, ACT 2601, Australia

FEATURES

source 1..5501
 /organism="Expression vector pGH/F2.2-163"
 /db_xref="taxon:56100"
 /note="pGEX-2T derivative"
 258..1475
 /gene="GST/F2.2-163"
 258..1475
 /gene="GST/F2.2-163"
 /codon_start=1
 /transl_table=11
 /product="GSTmF2a/2-163"
 /protein_id="AAC08715.1"
 /db_xref="GI:3002495"
 /translation="MSPILGYKRIKGIQVOPFLLEYLEEKYEELHEDDEKRRNK
 KRELLEPNLPDYITDGKRLQSMATIRITADKINMLGGCKEERAEISMEGAVLDI
 RGVSRIVYSKDEFTLKVDFLSKLPKMLKEDRLCHRTYINGDHVTHDPMYDALD

VVLVDPKCLDAEPKLVCFKRIEAIPOIDKYLKSKYIAMPLOGQWATFGGDDHPK
 SDLVKRGSPVSHHHHHHGEKILYODPGNFDPSRSSSGSPAHSEYSSGGGQQ
 KRIVMPGSGSAFIPTINAITTSODLQWQPTVITSMNPPRSHVPSPLGLASVP
 GMAHPKPVITIGTIVYGRRRRDQLSPHEEEKRIRRENRKLAACORNRRELTE
 KIOAETEELEEK"
 1909..2769
 /codon_start=1
 /transl_table=11
 /product="beta-lactamase"
 /protein_id="AAC08716.1"
 /db_xref="GI:3002497"
 /translation="MSIOHFRVALIPFAFLCPVPAHETLVKKVDAEDQIARVY
 IELDINSKILFESFRPEERFPMSTFEKVLICGAVLSRDVAGQDGLRIRHYSQNDLVE
 YSPVEKHLTDMTVRELCSAAITMSDNTAAKMLLTITIGPPELTAFLIMQSDVTRL
 DWPEPELNEAIIPNDERDPTMPAAMATLNLITGELLTGLTAHQOILDMHEAKVAGPL
 LRSALPAGWFIADKSGAGERSRGIIAALCPGKRISRIVIVITTSQATMDERNQIA
 ELGASLILKHW"
 3641..4932
 /codon_start=1
 /transl_table=11
 /product="lac operon repressor"
 /protein_id="AAC08717.1"
 /db_xref="GI:3002498"
 /translation="MVNKKPVLYDAEYAGVSYOTVSHVYNQASHVSATREKVEAA
 MAELVYIPNRVQDLAKQSLIGVATSSLAHAPSQIVAAIKSRADQIGASVSMV
 ERSQVCAKAAVHNLDAORVSLIINYPLDDDAIAVEAATNVPALFYLDQPIIN
 SLIFSEHEDCTRLGVHVALHQOIALAGPLSSVSARLGMKHYLRNIOIPAE
 RGDMSAMSGFOOTMOMNEGIVPTMAVLANOMALGANRATSEGLRPGADISVYG
 DTEBSSCTIPEPLTTIKODFRILGOTSDRLQLSOGQAVKRNQILPVLVRRKTTLA
 PMTGTASPPALDLSLMQARVSRLESQ"

BASE COUNT 1353 a 1367 c 1444 g 1337 t
 ORIGIN

Query Match 62.1%; Score 681; DB 12; Length 5501;
 Best Local Similarity 100.0%; Pred. No. 5e-165;

Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCCCTATACCTAGTATTTGGAATAATGAAGGCGCTTGCAACCCATCGACTTCTT 60
 Db 258 ATGCCCTATACCTAGTATTTGGAATAATGAAGGCGCTTGCAACCCATCGACTTCTT 317
 Qy 61 TTGGAATATCTTGAAGAAAATATGAAGACATTTGATGAGCGGATGAAGTGTATA 120
 Db 318 TTGGAATATCTTGAAGAAAATATGAAGACATTTGATGAGCGGATGAAGTGTATA 377
 Qy 121 TGCGCAACAAAAGTTGATTTGGGTTGGAGTTTCCCATCTTCTTATATTTGAT 180
 Db 378 TGCGCAACAAAAGTTGATTTGGGTTGGAGTTTCCCATCTTCTTATATTTGAT 437
 Qy 181 GGTGATGTTAAATTAACACAGCTATAGGCCATCATAGTATATAGCTGACAGCAAC 240
 Db 438 GGTGATGTTAAATTAACACAGCTATAGGCCATCATAGTATATAGCTGACAGCAAC 497
 Qy 241 ATGTTGGGTGTTGTCACAAAGAGCGTGCAGAGATTTCAATGCTTGAAGAGCGGTTTG 300
 Db 498 ATGTTGGGTGTTGTCACAAAGAGCGTGCAGAGATTTCAATGCTTGAAGAGCGGTTTG 557
 Qy 301 GATATTAGTACGCTGTTTTCGAGAAATTCATATAGTAAGACTTTGAACCTCAAGTT 360
 Db 558 GATATTAGTACGCTGTTTTCGAGAAATTCATATAGTAAGACTTTGAACCTCAAGTT 617
 Qy 361 GATTTCTTAGACAGCTACCTGAATGCGAATGTTGGAAGATGTTATGCTCAATAA 420
 Db 618 GATTTCTTAGACAGCTACCTGAATGCGAATGTTGGAAGATGTTATGCTCAATAA 677
 Qy 421 ACATATTTAAATGATGATCATGTAACCCATCCGACTTCATGTTGATGACGCTTGAT 480
 Db 678 ACATATTTAAATGATGATCATGTAACCCATCCGACTTCATGTTGATGACGCTTGAT 737
 Qy 481 GTTGTGTTTATACATGACCAATGTCCTGATGCGTTCACCAAAATTAAGTTGTTTAA 540
 Db 738 GTTGTGTTTATACATGACCAATGTCCTGATGCGTTCACCAAAATTAAGTTGTTTAA 797

QY 541 AACGATTTGAACTTATCCCAATTTGATTAAGTACTTGAATCCAGCATATATAGCA 600
DB 798 AACGATTTGAACTTATCCCAATTTGATTAAGTACTTGAATCCAGCATATATAGCA 857
QY 601 TGGCCCTTTCAGGCGTGGCAAGCAGCTTGGTGGTGGGCGACCATCTCCCAAAATCGCAT 660
DB 858 TGGCCCTTTCAGGCGTGGCAAGCAGCTTGGTGGTGGGCGACCATCTCCCAAAATCGCAT 917
QY 661 CTGGTTCGCGTGGATCCCG 681
DB 918 CTGGTTCGCGTGGATCCCG 938

RESULT 11
EV085202 5504 bp DNA circular SYN 31-MAR-1998
LOCUS EV085202 Expression vector pGH/F2.79-242, complete sequence.
DEFINITION U85202
ACCESSION U85202
VERSION U85202.1 GI:3002499
KEYWORDS
SOURCE Expression vector pGH/F2.79-242.
ORGANISM Expression vector pGH/F2.79-242
artificial sequence; vectors.
REFERENCE 1 (bases 1 to 5504)
AUTHORS Cahill, M.A.
TITLE Phosphorylation of Transcription Factors Elk-1 and Fra-2 by Associated Serum-dependent Kinase Activity
JOURNAL 2 (bases 1 to 5504)
REFERENCE Direct Submission
AUTHORS JCSMR, Australia National University, P.O. Box 334, Canberra, ACT
JOURNAL 2601, Australia

FEATURES
source
1. 5504
Location/Qualifiers
/organism="Expression vector pGH/F2.79-242"
/db_xref="taxon:56101"
/note="pGEX-2T derivative"
258. 1478
/gene="GST/F2.79-242"
258. 1478
/gene="GST/F2.79-242"
/codon_start=1
/transl_table=1
/product="GSTM2/79-242"
/protein_id="AAC08718.1"
/db_xref="GI:3002500"
/transl_table=1
/codon_start=1
/product="beta-lactamase"
/protein_id="AAC08720.1"
/db_xref="GI:3002502"
/transl_table=1
/codon_start=1
/product="MSI0HFRVALIPFAACLPVFAHPETLVKVKDAEDQLGARVY
IEDLNSGKILLESFRRPEERPMNSDFYLLCGAVLSVDAAGEOLGRIRHYSQNDLVE
VSPYRKHLTDGTVRELTSAITMSDNTAAILLTIGGPKETAFIHMNSGLKVEI
DWMPELNEAIPNDERDTTTPAAMATTIRKLILGLITLASROOLDIMRADKYAGPL
LRSAIPAGMPLADKSGAGERSKGTIALPDPGPKRIIVITVITGSGATMDERNRDLA
ELIGASLIKHW"
3841. 4935
/codon_start=1
/transl_table=1
/product="lac operon repressor"

CDS
CDS

/protein_id="AAC08719.1"
/db_xref="GI:3002501"
/translation="MVNKKPVTLLDVAEYAGSVQYTSRVYNQASHVSATREVEEA
MAELNYPNRAQOLAKGOSLLIVATSSIALHPRSQIVAAIKSRADLQASVSVK
ERSVEACKAIVNHLAIVRSGLIINPLDDQDAIVAEACTNVALEFLVSDQTPIN
SIFSHEDGTRFLGVSHVVLGHOQIALGDSVSRRLRAGWKYLRNOQPIAE
REGDMSASMGQOQTMOMLNEGIVPTAMLVANDQALAMRAITSGSRVAGDISVGY
DDTDSGCIYIPPLTIIKODFRLGOTSVDRLLQLSQQAVVGNQLLPLVSLKRTTILA
PNQOTASPRALDLSLMOLARQVSRLESGO"

BASE COUNT 1358 a 1354 c 1463 g 1329 t
ORIGIN

Query Match 62.1%; Score 681; DB 12; Length 5504;
Best Local Similarity 100.0%; Pred. No. 5e-165;
Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTCCCTTATACATAGCTTATTTGAATAATTAAGGCTTGTGCAACCCACTGACTCTT 60
DB 258 ATGTCCCTTATACATAGCTTATTTGAATAATTAAGGCTTGTGCAACCCACTGACTCTT 317
QY 61 TTGGAATATCTTGAAGAAATATGAAGACATTTGTATGAGCGGATGAGTGATATAA 120
DB 318 TTGGAATATCTTGAAGAAATATGAAGACATTTGTATGAGCGGATGAGTGATATAA 377
QY 121 TGGCGAACAACAAAGTTGGAATTTGGGTTGGAGTTGCCAATCTTCTATTATATATGAT 180
DB 378 TGGCGAACAACAAAGTTGGAATTTGGGTTGGAGTTGCCAATCTTCTATTATATATGAT 437
QY 181 GGTGATGTAAATTAACACAGCTATAGCCATCATACCTTATATAGCTGACACACAC 240
DB 438 GGTGATGTAAATTAACACAGCTATAGCCATCATACCTTATATAGCTGACACACAC 497
QY 241 ATGTGGGTGGTGTGTCACAAAGAGGCTGACAGATTTTAAGCTTGAAGAGCGGTTTG 300
DB 498 ATGTGGGTGGTGTGTCACAAAGAGGCTGACAGATTTTAAGCTTGAAGAGCGGTTTG 557
QY 301 GATATTTAGATACGGTGTGTCAGAAATTTGATATAGTAAGACTTTGAACTCTCAAGTT 360
DB 558 GATATTTAGATACGGTGTGTCAGAAATTTGATATAGTAAGACTTTGAACTCTCAAGTT 617
QY 361 GATTTTCTTACCAAGCTACCTGAAATGCTGAAATGTTTGAAGATCGTTATGTCATATA 420
DB 618 GATTTTCTTACCAAGCTACCTGAAATGCTGAAATGTTTGAAGATCGTTATGTCATATA 677
QY 421 ACATATTTAATTTGATGATCATGTAAACCATCTGATCATGTTGTATGAGCGTCTTGAT 480
DB 678 ACATATTTAATTTGATGATCATGTAAACCATCTGATCATGTTGTATGAGCGTCTTGAT 737
QY 481 GTTGTATTTATCATGAGACCCCAATGCTGATCGTGTCCCAAAATTAGTTGTTTTAAA 540
DB 738 GTTGTATTTATCATGAGACCCCAATGCTGATCGTGTCCCAAAATTAGTTGTTTTAAA 797
QY 541 AACGATTTGAACTTATCCCAATTTGATTAAGTACTTGAATCCAGCATATATAGCA 600
DB 798 AACGATTTGAACTTATCCCAATTTGATTAAGTACTTGAATCCAGCATATATAGCA 857
QY 601 TGGCCCTTTCAGGCGTGGCAAGCAGCTTGGTGGTGGGCGACCATCTCCCAAAATCGCAT 660
DB 858 TGGCCCTTTCAGGCGTGGCAAGCAGCTTGGTGGTGGGCGACCATCTCCCAAAATCGCAT 917
QY 661 CTGGTTCGCGTGGATCCCG 681
DB 918 CTGGTTCGCGTGGATCCCG 938

RESULT 12
EV085203
LOCUS EV085203 5516 bp DNA circular SYN 31-MAR-1998
DEFINITION EV085203 Expression vector pGH/F2.159-327, complete sequence.
ACCESSION U85203
VERSION U85203.1 GI:3002503
KEYWORDS

CDS

258..1712
 /gene="GST/F2.2-242"
 /codon_start=1
 /transl_table=11
 /product="GstmFra2/2-242"
 /protein_id="AAC08724.1"
 /db_xref="GI:3002508"
 /translation="MSPIIGYWKIKGLVQPTRLLEYLEEKYEHLERDEGDKWRNK
 KFELEGPENLPYIDGVKLQSMALIRYIADKRNMLGGCKREKAEISMLGAVLID
 RYGVSRIVASKDEFTLVKDFLSKLEPMLEKEDRCHKTLYNGDHVTHPDMLYALD
 VVLYMDPMLDAPEKLVCFKKRIEAIPOIDYIKSKXYIADPLOCWOMOTFGGDHPK
 SDIYVPGSPVEMSHHHHNGEKIRISPYIPGNETSSNGSSGSPAHASISYSGGGQQ
 KFRVDPSSGSAFIPTINAITSDQLMWVNPVITTSNSNPYPSHPYPLGLASVP
 GHMALPRGVKITGTTVGRRRRDEQLSPEEEKRIRRENRKLAALACRRRRELT
 KLOAETLEEEKSGLOKEIAELOEKELFPMKVAHGHPVCKISPEERSRPTSGLOS
 LRCGSGAVGPVAVQEPPEEDSPSSSAGMDKTR"

CDS

2146..3006
 /codon_start=1
 /transl_table=11
 /product="beta-lactamase"
 /protein_id="AAC08726.1"
 /db_xref="GI:3002510"
 /translation="MSIOHFVALIPPEAPACLVFAHPETLVKDAEDOLGARVQ
 IEIDNSGKILSEPRPERPFPMSTFKVLCGAVLSRDAGQEDGRIRHYSQNDLVE
 YSPYTERKLTIGMTRELCSAITMSDNTANLLITGPERELTAFLHMGDHYATL
 DRWPELNEALIPNEDRTMPAAMATTLRLKLTGELLTLAGRQOLIDMEADKVAPL
 LRSALPAGMFTADKSGAGENGSGIILALGPDKPSRIVITYTGSQATMDERNROIA
 EIGASLIKHW"
 4078..5169
 /codon_start=1
 /transl_table=11
 /product="lac operon repressor"
 /protein_id="AAC08725.1"
 /db_xref="GI:3002509"
 /translation="MVNVKPYTLVDVAEYAGVSYQTVSRVYNQASHVSAKREKVEEA
 MAELNIIPNRYAQLAGKOSLIGVATSTALHAPSOIVAAIKRAQLDGLASVYSAV
 ERSGVEACKAAYHNHLLAKORVSGLIINPLDDODIAYEACTIONPALFDVSDPTIN
 SIIRSHEDGTRLEVHVALGHQIALLAGLSSVARSRLAGHKYLTNRQDIPLE
 REGDWSASGSGQOTMOMINEGIPTAMLVANDOMALAMRAITSGSLRVGADISVGY
 DDTEDSSCYIPPLTTIKODPRLGOTGVDRILLOISQCAVAVGNQLPLVSIYVKRKTTLA
 PMTQTASPRALADSIMOLAROVSRLEFSQ"

BASE COUNT 1405 a 1438 c 1523 g 1372 t
 ORIGIN

Query Match 62.1%; Score 681; DB 12; Length 5738;
 Best Local Similarity 100.0%; Pred. No. 5e-165;
 Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCCCTACTAGTATTGGAATTAAGGCTTGTCACCACTGACTCTT 60
 DB 258 ATGCCCTACTAGTATTGGAATTAAGGCTTGTCACCACTGACTCTT 317
 QY 61 TTGGAATCTTGAAGAAATATGAGAGCATTTGTATGACGGGATGAAGTATAA 120
 DB 318 TTGGAATCTTGAAGAAATATGAGAGCATTTGTATGACGGGATGAAGTATAA 377
 QY 121 TGGCGAAACAAAATTTGAATGGTTTGCATTCCTTCTTATATATGAT 180
 DB 378 TGGCGAAACAAAATTTGAATGGTTTGCATTCCTTCTTATATATGAT 437
 QY 181 GGTGATGTTAATACACACTATGCGCATATAGCTTATATAGTGACACACAA 240
 DB 438 GGTGATGTTAATACACACTATGCGCATATAGCTTATATAGTGACACACAA 497
 QY 241 ATGTTGGTGTGTCACAAAGACGTGACAGATTTCAATGCTTGAAGGAGGCTTTG 300
 DB 498 ATGTTGGTGTGTCACAAAGACGTGACAGATTTCAATGCTTGAAGGAGGCTTTG 557
 QY 301 GATATTAGATCGGTGTTGAGAAATGACATATAGTAAGACTTTGAAACTCTCAAGTT 360
 DB 558 GATATTAGATCGGTGTTGAGAAATGACATATAGTAAGACTTTGAAACTCTCAAGTT 617

QY 361 GATTTCTTAGCAGCTACCTGAAATGCGAATGTTGGAAGATCGTTATATGATATAA 420
 DB 618 GATTTCTTAGCAGCTACCTGAAATGCGAATGTTGGAAGATCGTTATATGATATAA 677
 QY 421 ACATATTTAAATGTCATCATGTAACCCATCTGACTTCATGTTATGAGACCTTTGAT 480
 DB 678 ACATATTTAAATGTCATCATGTAACCCATCTGACTTCATGTTATGAGACCTTTGAT 737
 QY 481 GTTCTTTATACATGAGACCAATGTCCTGATGCGTTCCCAAAATAGTTGTTTAAA 540
 DB 738 GTTCTTTATACATGAGACCAATGTCCTGATGCGTTCCCAAAATAGTTGTTTAAA 797
 QY 541 AAACGATTTGAAGCTATCCCAATGTAAGTACTTGAATCCAGAGATATATACA 600
 DB 798 AAACGATTTGAAGCTATCCCAATGTAAGTACTTGAATCCAGAGATATATACA 857
 QY 601 TGGCCTTTGCAAGCGCTGGCAAGCCAGCTTTGGTGTGGCGACCATCTCCAAATCGAT 660
 DB 858 TGGCCTTTGCAAGCGCTGGCAAGCCAGCTTTGGTGTGGCGACCATCTCCAAATCGAT 917
 QY 661 CTGCTTCCCGCGTGATCCCCG 681
 DB 918 CTGCTTCCCGCGTGATCCCCG 938

RESULT 14
 EVU85205
 LOCUS 5756 bp DNA circular SYN 31-MAR-1998
 DEFINITION Expression vector pGH/F2.79-327, complete sequence.
 ACCESSION U85205
 VERSION 085205.1
 KEYWORDS GI:3002511
 SOURCE
 ORGANISM
 Expression vector pGH/F2.79-327.
 artificial sequence; vectors.
 1 (bases 1 to 5756)
 AUTHORS Cahill,M.A.
 TITLE Phosphorylation of Transcription Factors Elk-1 and Fra-2 by Associated Serum-dependent Kinase Activity
 JOURNAL Unpublished
 2 (bases 1 to 5756)
 AUTHORS Cahill,M.A.
 TITLE Direct Submission
 JOURNAL Submitted (14-JAN-1997) Division of Immunology & Cell Biology,
 JCSMR, Australia National University, P.O. Box 334, Canberra, ACT
 2601, Australia
 FEATURES
 source
 Location/Qualifiers
 1..5756
 /organism="Expression vector pGH/F2.79-327"
 /db_xref="taxon:56104"
 /note="pGEX-2T derivative"
 258..1730
 /gene="GST/F2.79-327"
 258..1730
 /gene="GST/F2.79-327"
 /codon_start=1
 /transl_table=11
 /product="GstmFra2/79-327"
 /protein_id="AAC08727.1"
 /db_xref="GI:3002512"
 /translation="MSPIIGYWKIKGLVQPTRLLEYLEEKYEHLERDEGDKWRNK
 KFELEGPENLPYIDGVKLQSMALIRYIADKRNMLGGCKREKAEISMLGAVLID
 RYGVSRIVASKDEFTLVKDFLSKLEPMLEKEDRCHKTLYNGDHVTHPDMLYALD
 VVLYMDPMLDAPEKLVCFKKRIEAIPOIDYIKSKXYIADPLOCWOMOTFGGDHPK
 SDIYVPGSPVEMSHHHHNGEKIRISPYIPGNETSSNGSSGSPAHASISYSGGGQQ
 KFRVDPSSGSAFIPTINAITSDQLMWVNPVITTSNSNPYPSHPYPLGLASVP
 GHMALPRGVKITGTTVGRRRRDEQLSPEEEKRIRRENRKLAALACRRRRELT
 KLOAETLEEEKSGLOKEIAELOEKELFPMKVAHGHPVCKISPEERSRPTSGLOS
 LRCGSGAVGPVAVQEPPEEDSPSSSAGMDKTR"
 2146..3024
 /codon_start=1
 /transl_table=11
 /product="beta-lactamase"

/protein_id="AAC08729.1"
/db_xref="GI:3002514"
/translation="MSIQHFRVALLPFAAFLPYFAHPETLVKVDADOLGARVGY
IELDLNSGKILSESPREPRMSTFKVLICGAVLSRVDAGEOLGRRIHNSONDLVE
YSPYTERKHLTDGMTVRELCSSAITSNDNTAANLLITTTGGPELTAFLNNKGDHVTRE
DRWEPELNEAIPNDRDITPMAMAATTIRKLITGLTLLASRQOITIDMEADKVAQPL
LRSAIPAGMFIADKSGAGERSRGITIAALGPDGKRSRIVVIYTTGSOATMDERNRQIA
EIGASLIKHW"
4096..5187
/codon_start=1
/product="lac operon repressor"
/protein_id="AAC08728.1"
/db_xref="GI:3002513"
/translation="MVNWKPYTLVDVAEYAGVSYQTVSRVNVQASVSAKREKVEEA
MAELNYIPNRVAQOLAGKOSLLIGVATSSLAHAPSOYAAIKSRADOLGASVYVM
ERSGVEACKAAVHNLAORVSGLIINYPDDDDAIVAEACNTNPALFLDVSDFPIN
SLIFSHEDGTRIGVEHLVALGHOQIALLAGPLSSASARLRAGMKRYLTRNDIQPIAE
REGDMSASGFOOTMOMNEGIVPTAMLVANDOMALGAMRAITGESLRVAGDISVYG
DDEDSQCIPLPLTIKODPFRLLGOTSVDRILQISOGQAVKGNOLLPSVLVKKRTTLLA
PNTQTSAPRALDSIMOLAROVSRLESQO"

BASE COUNT 1407 a 1448 c 1520 g 1381 t
ORIGIN

Query Match 62.1%; Score 681; DB 12; Length 5756;
Best Local Similarity 100.0%; Pred. No. 5e-165;
Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTCCCTTACTAGCTTATTGGAATAATTAAGGCGCTTGCGACCCACTGCATCTCT 60
DB 258 ATGTCCCTTACTAGCTTATTGGAATAATTAAGGCGCTTGCGACCCACTGCATCTCT 317
QY 61 TTGGAATATCTGGAAGAAAATTAAGAGCAATTTGATAGCGGCGATGAAGGTATAA 120
DB 318 TTGGAATATCTGGAAGAAAATTAAGAGCAATTTGATAGCGGCGATGAAGGTATAA 377
QY 121 TGGCGAACAACAAAGTTTGAATTTGGTTGAGTTTCCCAATCTCTTATTATTGAT 180
DB 378 TGGCGAACAACAAAGTTTGAATTTGGTTGAGTTTCCCAATCTCTTATTATTGAT 437
QY 181 GGTGATGTTAAATTAACACAGCTCTATAGCCATCATACGTTATATAGCTGACACAC 240
DB 438 GGTGATGTTAAATTAACACAGCTCTATAGCCATCATACGTTATATAGCTGACACAC 497
QY 241 ATGTTGGTGGTTGTCACAAAGAGCTGACAGATTTCATAGCTTGAGAGGCGTTTG 300
DB 498 ATGTTGGTGGTTGTCACAAAGAGCTGACAGATTTCATAGCTTGAGAGGCGTTTG 557
QY 301 GATATTAGATACGCTGTTTGCAGAAATTCATATAGTAAGACTTTGAACTCTCAAGTT 360
DB 558 GATATTAGATACGCTGTTTGCAGAAATTCATATAGTAAGACTTTGAACTCTCAAGTT 617
QY 361 GATTTTCTTACAGAGCTACCTGAATGCTGAANAATTTGAGAGATGTTTATGTCATAA 420
DB 618 GATTTTCTTACAGAGCTACCTGAATGCTGAANAATTTGAGAGATGTTTATGTCATAA 677
QY 421 ACATATTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
DB 678 ACATATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 737
QY 481 GTTGTGTTTATACATGAGCAACATGTCCTGATGCTGCTCCCAAAATTAAGTTTAA 540
DB 738 GTTGTGTTTATACATGAGCAACATGTCCTGATGCTGCTCCCAAAATTAAGTTTAA 797
QY 541 AAACGATTGAGAGCTATCCACAAATTTGATAGTACTGAAATCCGACAGATATATGCA 600
DB 798 AAACGATTGAGAGCTATCCACAAATTTGATAGTACTGAAATCCGACAGATATATGCA 857
QY 601 TGGCCTTTTGCAGGCTGCGACAGCAGCTTGTGTGGGACCATCCTCCAAATGCGAT 660
DB 858 TGGCCTTTTGCAGGCTGCGACAGCAGCTTGTGTGTGGGACCATCCTCCAAATGCGAT 917

QY 661 CTGCTTCCCGCTGAGATCCCG 681
DB 918 CTGCTTCCCGCTGAGATCCCG 938

RESULT 15
EV085206
LOCUS EV085206 5990 bp DNA circular SYN 31-MAR-1998
DEFINITION Expression vector pGH/F2.2-327, complete sequence.
ACCESSION U85206
VERSION U85206.1 GI:3002515
KEYWORDS
SOURCE ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS
TITLE
JOURNAL

Expression vector pGH/F2.2-327.
Expression vector pGH/F2.2-327
artificial sequence; vectors.
1 (bases 1 to 5990)
Cahill, M.A.
Phosphorylation of Transcription Factors Elk-1 and Fra-2 by
Associated Serum-dependent Kinase Activity
Unpublished
2 (bases 1 to 5990)
Cahill, M.A.
Direct Submission
Submitted (14-JAN-1997) Division of Immunology & Cell Biology,
JCSMR, Australia National University, P.O. Box 334, Canberra, ACT
2601, Australia
FEATURES
source
1..5990
/organism="Expression vector pGH/F2.2-327"
/db_xref="taxon:56105"
/note="pGEX-2T derivative"
258..1964
/gene="GST/F2.2-327"
258..1964
/gene="GST/F2.2-327"
/codon_start=1
/transl_table=1
/product="GSTmF2a2/2-327"
/protein_id="AAC08730.1"
/db_xref="GI:3002516"
/translation="MSPIIGYKIKIGLVOPTRLLLEYLEKEIEHLYEDDEDKMRK
KRELGEFENLPYIDGVKLRQSMATIRYIADKHNMLGCKERAREISMLGAVLID
RIGVSRISAKDFEPLKVDLSLKEPMLKMFEDRLCHKYVLDGHTHDFMLYALD
VVLVMDPMLDAFPLKVCFKRLEAIPOIDKYLKSKYIAMPLOGQATFFGGDHPK
SDLVPRGSPVENHHHHHGEKILYODYPGNEDESSRSRSGSPAAESYSGGGGQ
KRVMDPGSGSAFPIPTINAITSODLQWVOPVITSMENPSPRSPYPLDGLASVP
GHMALPSPGVITITGTVRRRDEQLSPEEEKRIRREKRLAAKCRNRRELTE
KLQAEFELEERKSGLOKEIAELQEKERLEFMKVAHGVCKISPEERSPTGLQS
LRGTSAVGPVVVKQEPREDSPSSAGMDKQRSYIKPISJAGGFGVEEPLHPPIV
VSTPAITGTGSLNLFVTPNVLEOESPSSSCSAHRSSSSGDSINSPTLL
AL"

CDS

/codon_start=1
/transl_table=11
/product="beta-lactamase"
/protein_id="AAC08732.1"
/db_xref="GI:3002518"
/translation="MSIQHFRVALLPFAAFLPYFAHPETLVKVDADOLGARVGY
IELDLNSGKILSESPREPRMSTFKVLICGAVLSRVDAGEOLGRRIHNSONDLVE
YSPYTERKHLTDGMTVRELCSSAITSNDNTAANLLITTTGGPELTAFLNNKGDHVTRE
DRWEPELNEAIPNDRDITPMAMAATTIRKLITGLTLLASRQOITIDMEADKVAQPL
LRSAIPAGMFIADKSGAGERSRGITIAALGPDGKRSRIVVIYTTGSOATMDERNRQIA
EIGASLIKHW"
4330..5421
/codon_start=1
/transl_table=11
/product="lac operon repressor"
/protein_id="AAC08731.1"

CDS

/codon_start=1
/transl_table=11
/product="lac operon repressor"
/protein_id="AAC08731.1"
/db_xref="GI:3002517"
/translation="MVNWKPYTLVDVAEYAGVSYQTVSRVNVQASVSAKREKVEEA
MAELNYIPNRVAQOLAGKOSLLIGVATSSLAHAPSOYAAIKSRADOLGASVYVM
ERSGVEACKAAVHNLAORVSGLIINYPDDDDAIVAEACNTNPALFLDVSDFPIN
SLIFSHEDGTRIGVEHLVALGHOQIALLAGPLSSASARLRAGMKRYLTRNDIQPIAE

REGDWSAMSGFOQTMOMINEGIVPTAMLVANDOMALGAMRAITTESGLRYGADISVVG
 DDEEDSSCYIPPLITIKODFRLIGQSVDRILQLSOGQAVKGNQLLPVSLVKRKTLLA
 PNCOTASPALADSIMOLAROVSRLESQO"
 BASE COUNT 1454 a 1532 c 1580 g 1424 t
 ORIGIN

Query Match 62.1%; Score 681; DB 12; Length 5990;
 Best Local Similarity 100.0%; Pred. No. 5e-165;
 Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGTCCCTATACCTAGGTTATTTGAAATTTAGGGCTTGTGCAACCCACTGACTTCTT	60
DB	258	ATGTCCCTATACCTAGGTTATTTGAAATTTAGGGCTTGTGCAACCCACTGACTTCTT	317
QY	61	TTGGAATATCTTGAAGAAAAATATGAGACATTTGATGAGCGGATGAGGTATATA	120
DB	318	TTGGAATATCTTGAAGAAAAATATGAGACATTTGATGAGCGGATGAGGTATATA	377
QY	121	TGCGGAACAAAAAGTTGAATGGGTTTGAGTTTCCCAATCTTCTTATATATGAT	180
DB	378	TGCGGAACAAAAAGTTGAATGGGTTTGAGTTTCCCAATCTTCTTATATATGAT	437
QY	181	GGTATGTAAATTAACACAGCTATGCGCATCATACGTTATATAGCTGACAGACAAC	240
DB	438	GGTATGTAAATTAACACAGCTATGCGCATCATACGTTATATAGCTGACAGACAAC	497
QY	241	ATGTTGGTGTGTTGCCAAGAGCGTGACAGATTTCATGCTTGAAGAGCGGTTTG	300
DB	498	ATGTTGGTGTGTTGCCAAGAGCGTGACAGATTTCATGCTTGAAGAGCGGTTTG	557
QY	301	GATATTAGATACGGTGTGGAATGCAATATAGTAAGACTTTGAACCTCTCAAGTT	360
DB	558	GATATTAGATACGGTGTGGAATGCAATATAGTAAGACTTTGAACCTCTCAAGTT	617
QY	361	GATTTTCTAGCAGCTACCTGAATGCGAATGTCGAAGATCGTTATATGATATAA	420
DB	618	GATTTTCTAGCAGCTACCTGAATGCGAATGTCGAAGATCGTTATATGATATAA	677
QY	421	ACATATTTAAATGTGATCATGTAAACCATCCTGACTTCATGTGTATGACGCTTGAT	480
DB	678	ACATATTTAAATGTGATCATGTAAACCATCCTGACTTCATGTGTATGACGCTTGAT	737
QY	481	GTTGTTTATACATGACCCCAATGTGCGTGATGCGTTCCCAAAATAGTTGTTTAA	540
DB	738	GTTGTTTATACATGACCCCAATGTGCGTGATGCGTTCCCAAAATAGTTGTTTAA	797
QY	541	AAACGATTTGAAGCTATCCCAATGTATAGTACTTGAATCCAGCAAGTATATAGCA	600
DB	798	AAACGATTTGAAGCTATCCCAATGTATAGTACTTGAATCCAGCAAGTATATAGCA	857
QY	601	TGGCCTTTGCAAGGGCTGGCAAGCCAGTTTGGTGGGCAACATCTCCAAATCGGAT	660
DB	858	TGGCCTTTGCAAGGGCTGGCAAGCCAGTTTGGTGGGCAACATCTCCAAATCGGAT	917
QY	661	CTGGTTCCGCGGTGATCCCG 681	
DB	918	CTGGTTCCGCGGTGATCCCG 938	

Search completed: March 18, 2001, 04:57:19
 Job time: 4243 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 18, 2001, 04:17:41 ; Search time 83.44 Seconds
(without alignments)
4934.400 Million cell updates/sec

Title: US-09-402-488A-1

Perfect score: 1096
Sequence: 1 ATGTCCCTACTAGAGTTA.....ATCCACTAGTTAGACCGG 1096

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 480022 seqs, 187831343 residues

Total number of hits satisfying chosen parameters: 960044

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
N.Geneseq_36: *
1: /cgn2_2/gcgdata/geneseq/geneseqn/NA1980.DAT: *
2: /cgn2_2/gcgdata/geneseq/geneseqn/NA1981.DAT: *
3: /cgn2_2/gcgdata/geneseq/geneseqn/NA1982.DAT: *
4: /cgn2_2/gcgdata/geneseq/geneseqn/NA1983.DAT: *
5: /cgn2_2/gcgdata/geneseq/geneseqn/NA1984.DAT: *
6: /cgn2_2/gcgdata/geneseq/geneseqn/NA1985.DAT: *
7: /cgn2_2/gcgdata/geneseq/geneseqn/NA1986.DAT: *
8: /cgn2_2/gcgdata/geneseq/geneseqn/NA1987.DAT: *
9: /cgn2_2/gcgdata/geneseq/geneseqn/NA1988.DAT: *
10: /cgn2_2/gcgdata/geneseq/geneseqn/NA1989.DAT: *
11: /cgn2_2/gcgdata/geneseq/geneseqn/NA1990.DAT: *
12: /cgn2_2/gcgdata/geneseq/geneseqn/NA1991.DAT: *
13: /cgn2_2/gcgdata/geneseq/geneseqn/NA1992.DAT: *
14: /cgn2_2/gcgdata/geneseq/geneseqn/NA1993.DAT: *
15: /cgn2_2/gcgdata/geneseq/geneseqn/NA1994.DAT: *
16: /cgn2_2/gcgdata/geneseq/geneseqn/NA1995.DAT: *
17: /cgn2_2/gcgdata/geneseq/geneseqn/NA1996.DAT: *
18: /cgn2_2/gcgdata/geneseq/geneseqn/NA1997.DAT: *
19: /cgn2_2/gcgdata/geneseq/geneseqn/NA1998.DAT: *
20: /cgn2_2/gcgdata/geneseq/geneseqn/NA1999.DAT: *
21: /cgn2_2/gcgdata/geneseq/geneseqn/NA2000.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1096	100.0	1096	20	V83966 DNA encoding the f
2	685	62.5	720	20	X01312 GST coding sequenc
3	684.2	62.4	3591	21	Z58406 DNA encoding funon
4	684.2	62.4	3591	21	Z60645 DNA encoding a fus
5	684	62.4	2475	21	Z32916 Human HCP/GST fusi
6	683.4	62.4	1866	21	Z32917 Human HCP phosphod
7	683.2	62.3	2079	21	Z58400 DNA encoding GST-t
8	683.2	62.3	2079	21	Z58407 DNA encoding an am
9	683	62.3	2490	21	Z58407 DNA encoding GST-t
10	683	62.3	2490	21	Z60646 DNA encoding a fus
11	682.2	62.2	3618	21	Z58405 DNA encoding funon
12	682.2	62.2	3618	21	Z60644 DNA encoding a fus

13	681	62.1	699	16	O86774 Glutathione-s-tran
14	681	62.1	1695	16	T05329 DNA encoding TCR27
15	681	62.1	1932	16	T05332 DNA encoding TCR27
16	681	62.1	3175	20	Z20066 DNA encoding gluta
17	681	62.1	3867	20	Z20067 DNA encoding gluta
18	679.4	62.0	693	16	O87508 Glutathione-s-tran
19	678.4	61.9	1281	16	T03457 Turkey prolactin/G
20	678.2	61.9	1740	21	Z58065 CAP-A-B fusion pro
21	678	61.9	816	13	O22843 DNA encoding TCR27
22	677.8	61.8	1065	16	T05330 DNA encoding TCR27
23	677	61.8	1119	18	T90496 GST-Treponea pall
24	675.2	61.6	1335	13	O22855 pGEX-2T-15:17 enco
25	674	61.5	759	13	O22840 pGEX-2T-CAP-A enco
26	674	61.5	3423	16	T00771 GST-SBP Fusion gen
27	673.6	61.5	1194	16	O86776 GST-SBP-1 gene fus
28	673.2	61.4	759	13	O22841 pGEX-2T-CAP-B enco
29	673.2	61.4	759	13	O22842 pGEX-2T-CAP-C enco
30	667.4	60.9	897	16	O82895 Sequence encoding
31	663.4	60.5	751	19	V61099 Plasmid pGEX-PH 75
32	663.4	60.5	1128	16	O95552 GST-calmodulin fus
33	663.4	60.5	1239	18	T93979 DNA for oligopepti
34	663.4	60.5	1419	16	T05333 DNA encoding TCR27
35	663.4	60.5	1830	19	V61100 Plasmid pGEX-PH 18
36	663.4	60.5	1820	20	X77118 GS-GC6 fusion prot
37	662.8	60.5	2095	20	V55614 pGEX-MSP-1 methylas
38	661.8	60.4	951	17	T30371 pGEX/hp57 plasmid
39	661.8	60.4	1842	17	T30370 GST/truncated huma
40	661.8	60.4	2067	17	T30369 GST/human p57 fusi
41	661	60.3	924	16	T05331 DNA encoding TCR27
42	660.4	60.3	1851	19	V13177 GST-cyclin E-PEST*
43	659.6	60.2	894	13	O22852 pGEX-3X-693:691 en
44	658.6	60.1	858	13	O22853 pGEX-3X-15:18 enco
45	658.6	60.1	948	13	O22839 pGEX-3X-690:694 en

ALIGNMENTS

RESULT 1	
ID	V83966
DB	V83966 standard; DNA; 1096 BP.
AC	V83966;
XX	
DT	09-MAR-1999 (first entry)
DE	DNA encoding the fusion protein GST-chymosin pro-peptide-Hirudin.
XX	
KW	Fusion protein; bovine chymosin pro-peptide; leech; hirudin;
KW	anticoagulant protein; autocatalytically maturing zymogen;
KW	carp growth hormone; vaccine; ss.
XX	
OS	Synthetic.
OS	Bos sp.
OS	Hirudo medicinalis.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	1..1032
FT	/tag= a
FT	/product= GST-chymosin.pro-peptide-hirudin
FT	1..708
FT	/tag= b
FT	/note= "encodes GST"
FT	709..834
FT	/tag= c
FT	/note= "encodes bovine chymosin pro-peptide"
FT	835..1029
FT	/tag= d
FT	/note= "encodes leech hirudin; sequence has been
XX	optimised for plant codon usage"
XX	W09849326-A1.

PD 05-NOV-1998.
 XX 23-APR-1998; 98WO-CA00398.
 PF 25-APR-1997; 97US-0044254.
 PR
 XX (SEMB-) SEMBIOSYS GENETICS INC.
 PA
 PI Alcantara J, Moloney M, Van Rooijen G;
 XX WPI: 1999-059646/05.
 DR P-PSDB; W87636.
 XX
 PT Preparation of recombinant polypeptides - by expression of a fusion
 PT protein comprising a pro-peptide derived from an autocatalytically
 PT maturing zymogen and a heterologous polypeptide
 XX
 PS Claim 27; Fig 1; 44pp; English.
 XX
 CC The present sequence encodes a fusion protein comprising
 CC glutathione-S-transferase (GST)-bovine chymosin pro-peptide-leech
 CC hirudin. The chymosin pro-peptide sequence is placed upstream of the DNA
 CC sequence encoding the leech anticoagulant protein hirudin. The fusion
 CC protein was made to exemplify the invention. The specification describes
 CC a method for preparing a recombinant polypeptide in a host cell. A
 CC chimeric nucleic acid sequence encoding a fusion protein comprising
 CC a pro-peptide derived from an autocatalytically maturing zymogen linked
 CC a protein heterologous to the pro-peptide, is introduced into the host
 CC cell. The host cells are then grown to produce the fusion protein.
 CC Altering the environment of the fusion protein allows cleavage of the
 CC pro-peptide from the fusion protein to release the recombinant
 CC polypeptide. The method can be used for the preparation of recombinant
 CC polypeptides such as hirudin or carp growth hormone. The fusion
 CC proteins can be used for delivering to a human or animal a therapeutic
 CC or nutritional polypeptide such as a vaccine, a peptide antibiotic, a
 CC cattle feed enzyme, a cytokine, a gastric lipase or a lactase.
 XX
 SQ Sequence 1096 BP; 298 A; 230 C; 271 G; 297 T; 0 other;

Query Match 100.0%; Score 1096; DB 20; Length 1096;
 Best Local Similarity 100.0%; Pred. No. 3.8e-292;
 Matches 1096; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCCCTATCTAGTATGAAATTAAGGCGCTTGCAACCCAGCTGCTCT 60
 DB 1 atgcccctatactctgattatgaaatlaaaggccttggcaaccactgacttctt 60
 QY 61 TTGAATATCTTGAAGAAATATGAGAGCATTTGTATGAGCGGATGAGGTATAA 120
 DB 61 ttggaatatcttgaagaaatatagaagacatttgalagcgagatgaaggtatataa 120
 QY 121 TGGCAAAACAAAAGTTGGAATGGGTTGGAGTTCCCAATCTCTTATATATGAT 180
 DB 121 tggcgaaacaaaagttggaattgggttggagttcccaactcttctattatgat 180
 QY 181 GGTGATGTTAATTAACACAGTCTATGGCCATCATAGCTATATAGCGACACAAC 240
 DB 181 ggtgatgttaatttaacacagctctatggccatcatacgttatatacgcgacaagacaac 240
 QY 241 ATGTTGGTGTGTTGCCAAAAGAGCGTCGACAGATTTCAATGCTTGAAGAGCGGTTTGG 300
 DB 241 atgttgggtgttggtgccaaaagagcgctcgagagatttcaatgcttgaagagcggttttg 300
 QY 301 GATATTAGATAGGTTTCCGAGATTTGATATAGTAAGAGTTTGAACCTCTCAAGTT 360
 DB 301 gatattagataggtttccgagatTTGATATAGTAAGAGTTTGAACCTCTCAAGTT 360
 QY 361 GATTTTCTAGCAAGTACTCAAAATGCGAAATGTTGAGAGATGTTATGATATAA 420
 DB 361 gattttcttagcaagtactcaaaatgctgaaatgTTGAGAGATGTTATGATATAA 420
 QY 421 ACATATTTAATGTGATCATGTAACCCATCTGACTTCATGTTGATGACGCTTGAT 480

DB 421 acatattaaatgtgtatcatgtaaccacatccctgactcatgttattgtagcgtctgat 480
 QY 481 GTTGTTTTATCATGAGACCCCAATGCTCGATGCTTCCCAAAATAGTTTATTA 540
 DB 481 gtgtttttatcatgagaccccaatgcttgccttgatgtcgttcccaaatagttgtttaa 540
 QY 541 AAAGTATTTGAAGCTATCCCAAAATGATAGTACTTGAATCAGACAGTATATAGCA 600
 DB 541 aaagtatttgaagctatcccaaaatgattagatagtaacttgaatcagagatataagca 600
 QY 601 TGGCTTTTGCAGGGCTGAGCAACGCTTGTGTGTGCGGACATCTCCAAATTCGAT 660
 DB 601 tggccttttgcagggtgagcaacgcttgtgtgtgctgacacatctccaaatcggat 660
 QY 661 CTGTTTCCGCTGATATCCCAATTCGCGGTGATGACGACGCGGCGGTGATATCACC 720
 DB 661 ctgtttccgcgtgatatcccaatctcgggtgactgacgagcgcgctgagataacc 720
 QY 721 AGGATCCCTCTGTACAAAGCAAGCTCTGTAGAGAGGCGCTGAGAGAGCATGGCTTCG 780
 DB 721 aggatccctctgtacaaagcaagctctgtagagagcgctgagagcatggcttctcg 780
 QY 781 GAGGACTTCTGCAAGAACACAGCATATGCAATGACAGACAAATCTCCGCTTCGTCG 840
 DB 781 gaggaacttctgcagaaacacagcatatgcatagcagcaagtaactcggcttcgctgc 840
 QY 841 TATACCGACTTACCGAGTCCGCTGCTGCTGCTGCTGAGGCTTCCAAAGCTGCT 900
 DB 841 tataccgacttaccgagttccgctgctgctgctgctgctgctgctgctgctgctgct 900
 QY 901 GGTGAGGCTAACAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 DB 901 ggtgaggcttaacagtatctgctgctgctgctgctgctgctgctgctgctgctgct 960
 QY 961 GAGGAGACCCCAAGACCACTCCACAGACGAGTCTTGGAGAGATGCCAGAGAG 1020
 DB 961 gaggagaccccaagaccactccacagacgagtgcttggagagatccacagagag 1020
 QY 1021 TATGTCGATTAAGATCTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
 DB 1021 tatgtcगतtaagatctaatctgctgctgctgctgctgctgctgctgctgctgctgct 1080
 QY 1081 ACTAGTTCTAGAGCGG 1096
 DB 1081 actagttctagagcgg 1096

RESULT 2
 X01312
 ID X01312 standard; cDNA to mRNA; 720 BP.
 XX
 AC X01312:
 XX
 DT 13-APR-1999 (first entry)
 XX
 DE GST coding sequence.
 XX
 KW Cell surface nonexpressive functional protein; cell-permeating protein;
 KW protein screening; protein production; GST; glutathione-S-transferase;
 KW ds.
 OS Schistosoma japonicum.
 XX
 PN JP11009278-A.
 PD 19-JAN-1999.
 XX
 PD 23-JUN-1997; 97JP-0165788.
 PF 23-JUN-1997; 97JP-0165788.
 PR
 XX (SOME) SUMITOMO ELECTRIC IND CO.

XX WPI: 1999-145895/13.
 DR P-PSDB: W73909.
 XX
 PT Microbe having a nonexpressive functional protein on its surface -
 PT for screening and large-scale preparation of a selected function
 PT protein
 XX
 PS Disclosure: Page 9; 13pp; Japanese.
 XX
 CC This sequence encodes the glutathione-S-transferase (GST) protein, which
 CC was expressed in the microbe of the invention. The microbe has a cell
 CC surface nonexpressive functional protein on its surface, fused to a
 CC support protein derived from a cell-permeating protein. The microbe can
 CC be used for screening for biological activity of a cell surface
 CC nonexpressive protein. It can also be used for preparation of a selected
 CC functional protein in a large amounts.
 CC
 SQ Sequence 720 BP; 209 A; 130 C; 164 G; 217 T; 0 other;

Query Match 62.5%; Score 685; DB 20; Length 720;
 Best Local Similarity 98.1%; Pred. No. 4.2e-179;
 Matches 705; Conservative 0; Mismatches 10; Indels 4; Gaps 1;

QY 1 ATGTCCTCTATCTAGGTTATTGGAATTAAGGCGCTTGCAACCCACGCTGCTT 60
 DB 1 ATGTCCTCTATCTAGGTTATTGGAATTAAGGCGCTTGCAACCCACGCTGCTT 60
 QY 61 TTGGATATCTTGACAAAATATGAGAGCATTTGTATGAGCGCGATGAGTGATAA 120
 DB 61 TTGGATATCTTGACAAAATATGAGAGCATTTGTATGAGCGCGATGAGTGATAA 120
 QY 121 TGGCGAAGCAAAAGTTGATGAGTTGGATTTGCCATCTTCTTATTTATTTGAT 180
 DB 121 TGGCGAAGCAAAAGTTGATGAGTTGGATTTGCCATCTTCTTATTTATTTGAT 180
 QY 121 TGGCGAAGCAAAAGTTGATGAGTTGGATTTGCCATCTTCTTATTTATTTGAT 180
 DB 121 TGGCGAAGCAAAAGTTGATGAGTTGGATTTGCCATCTTCTTATTTATTTGAT 180
 QY 181 GGTGATGTTAATTAACACAGTCTATGCGCATCATGTTATATAGCTGACAGCAAC 240
 DB 181 GGTGATGTTAATTAACACAGTCTATGCGCATCATGTTATATAGCTGACAGCAAC 240
 QY 241 ATGTTGGGTGGTGTCCAAAAGCGCTGACAGATTTGATGAGAGCGGTTTGTG 300
 DB 241 ATGTTGGGTGGTGTCCAAAAGCGCTGACAGATTTGATGAGAGCGGTTTGTG 300
 QY 241 ATGTTGGGTGGTGTCCAAAAGCGCTGACAGATTTGATGAGAGCGGTTTGTG 300
 DB 241 ATGTTGGGTGGTGTCCAAAAGCGCTGACAGATTTGATGAGAGCGGTTTGTG 300
 QY 301 GATATTAGATACGCTGTTTCGACATTTGATAGTAAAGCTTGAACCTCCAAAGTT 360
 DB 301 GATATTAGATACGCTGTTTCGACATTTGATAGTAAAGCTTGAACCTCCAAAGTT 360
 QY 301 GATATTAGATACGCTGTTTCGACATTTGATAGTAAAGCTTGAACCTCCAAAGTT 360
 DB 301 GATATTAGATACGCTGTTTCGACATTTGATAGTAAAGCTTGAACCTCCAAAGTT 360
 QY 361 GATTTTCTTAGCAAGCTACCTGAATGCTGAANAATGTTGAGAGTGGTTATGATAA 420
 DB 361 GATTTTCTTAGCAAGCTACCTGAATGCTGAANAATGTTGAGAGTGGTTATGATAA 420
 QY 361 GATTTTCTTAGCAAGCTACCTGAATGCTGAANAATGTTGAGAGTGGTTATGATAA 420
 DB 361 GATTTTCTTAGCAAGCTACCTGAATGCTGAANAATGTTGAGAGTGGTTATGATAA 420
 QY 421 ACATATTAAATGATGATATGATACCATCCGACTTATGTTGATGACGCTTTGAT 480
 DB 421 ACATATTAAATGATGATATGATACCATCCGACTTATGTTGATGACGCTTTGAT 480
 QY 421 ACATATTAAATGATGATATGATACCATCCGACTTATGTTGATGACGCTTTGAT 480
 DB 421 ACATATTAAATGATGATATGATACCATCCGACTTATGTTGATGACGCTTTGAT 480
 QY 481 GTTGTATTAAATGATGATATGATACCATCCGACTTATGTTGATGACGCTTTGAT 540
 DB 481 GTTGTATTAAATGATGATATGATACCATCCGACTTATGTTGATGACGCTTTGAT 540
 QY 481 GTTGTATTAAATGATGATATGATACCATCCGACTTATGTTGATGACGCTTTGAT 540
 DB 481 GTTGTATTAAATGATGATATGATACCATCCGACTTATGTTGATGACGCTTTGAT 540
 QY 541 AAAGTATTGAAGCTATCCCAAAATGATAGTAAATCCAGCAAGTATATAGCA 600
 DB 541 AAAGTATTGAAGCTATCCCAAAATGATAGTAAATCCAGCAAGTATATAGCA 600
 QY 541 AAAGTATTGAAGCTATCCCAAAATGATAGTAAATCCAGCAAGTATATAGCA 600
 DB 541 AAAGTATTGAAGCTATCCCAAAATGATAGTAAATCCAGCAAGTATATAGCA 600
 QY 601 TGGCCCTTTGAGGCTGCGACAGCAGTTGGTGGGCGACCATCTCCAAAATGGAT 660
 DB 601 TGGCCCTTTGAGGCTGCGACAGCAGTTGGTGGGCGACCATCTCCAAAATGGAT 660
 QY 601 TGGCCCTTTGAGGCTGCGACAGCAGTTGGTGGGCGACCATCTCCAAAATGGAT 660
 DB 601 TGGCCCTTTGAGGCTGCGACAGCAGTTGGTGGGCGACCATCTCCAAAATGGAT 660
 QY 661 CTGTTGTCGCGGT----GATCCCGCAATTCGCGGTGATGAGCGCGCGCTCGGA 715
 DB 661 CTGTTGTCGCGGT----GATCCCGCAATTCGCGGTGATGAGCGCGCGCTCGGA 715

RESULT 3
 ID 258406
 XX 258406 standard; CDNA: 3591 BP.
 AC 258406;
 XX
 DT 23-MAY-2000 (first entry)
 DE
 XX DNA encoding fumonisin esterase-amino polyol amine oxidase fusion.
 KW Amino polyol amine oxidase; APAO; fumonisin esterase; mycotoxin;
 KW transgenic plant; detoxification; animal feed; silage;
 KW glutathione S transferase; alpha-amylase; signal peptide; maize;
 KW gs.
 XX
 OS Chimeric - Schistosoma japonicum.
 OS Chimeric - Bacterium.
 OS Chimeric - Exophiala spinifera.
 XX
 FH Key Location/Qualifiers
 FT sig-peptide 1..687
 FT /tag- a
 FT /product- "GST - polylinker"
 FT mat-peptide 677..3588
 FT /tag- b
 FT /product- "fumonisin-esterase-K:trAPAO fusion"
 FT misc-feature 2200..2202
 FT /tag- c
 FT /note- "extra lysine codon"
 XX
 PN WO200004159-A1.
 XX
 PD 27-JAN-2000.
 XX
 PF 08-JUL-1999; 99WO-US15454.
 XX
 PR 15-JUL-1998; 98US-0092936.
 PR 21-MAY-1999; 99US-0135391.
 XX
 PA (PION-) PIONEER HI-BRED INT INC.
 PA (CURA-) CURAGEN CORP.
 PI
 XX
 DU Duvick JP, Gilliam JT, Maddox JR;
 DR WPI: 2000-182425/16.
 DR P-PSDB: Y58917.
 XX
 PT New isolated polynucleotides, polypeptides useful for detecting and
 PT degrading fumonisin or structurally related mycotoxin in processed
 PT grain or in silage
 XX
 PS Example 13; Page 124-128; 154pp; English.
 XX
 CC The present sequence is that of a polynucleotide encoding a
 CC fusion protein (see Y58917) composed of glutathione S transferase
 CC and a bacterial fumonisin esterase (BSEI) joined via a peptide
 CC linker to K:trAPAO (see Y58909), a truncated, but still functional,
 CC amino polyol amine oxidase of Exophiala spinifera. The construct
 CC was designed for expression in Escherichia coli of a hybrid
 CC protein having BSEI and APAO activity. The invention provides APAO
 CC polynucleotides (see 258383-87) and polypeptides (see Y58900-05) of
 CC E. spinifera and Rhinocladia atroviens. The polynucleotides
 CC are used to transform plant cells normally susceptible to Fusarium
 CC or other toxin-producing fungus infection. Also provided are
 CC methods for expressing APAO, optionally as a fusion protein with
 CC fumonisin esterase, in transgenic plants, prokaryotic and non-plant
 CC eukaryotic systems. Methods for detoxification of grain, grain
 CC processing, silage, food crops and in animal feed and rumen
 CC microorganisms are also disclosed.
 XX
 SQ Sequence 3591 BP; 758 A; 1017 C; 1064 G; 752 T; 0 other;

Query Match 62.4%; Score 684.2; DB 21; Length 3591;
 Best Local Similarity 97.5%; Pred. No. 1.4e-178;
 Matches 695; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

```

OY 1 ATGTCCCTATACAGTATTGAAATTAAGGCGCTTGTCACACCCAGCTGCTT 60
Db 1 atgtccctatactagttatgtgaataatgaaggcctgtgcacccctgactctt 60
OY 61 TTGGAATATCTTGAGAAATAATGAGAGCATTTGTATGACGGATGAGAGTATAA 120
Db 61 ttggaatattcttgaagaataaataagaagcatttgtatgagcgcgaatgaatataa 120
OY 121 TCGCAAAACAAAAGTTTGAATTTGGTTTGGAGTTTCCCAATCTTCCTATTATTGAT 180
Db 121 tgcgaacaaaagtttgaatttggatttgggtttcccaatcttccattattatgtat 180
OY 181 GGTGATGTTAAATTAAACACAGCTCTATGGCCATCATACGTTATATAGCTGACACACAA 240
Db 181 ggtgatgttaaatataacacagctctatggccatcatatagcttgcacagcacac 240
OY 241 ATGTGGGAGGTGTTGCCAAAAGAGCGTGAAGATTTCATGCTTGAGGAGCGGTTTG 300
Db 241 atgtgggaggttggccaaaagagcggtgagagatttcaatgcttgaagagcggtttcg 300
OY 301 GATATTAGATACGCTGTTTCGAGAAATTCATATACATAAAGACTTTGAAACTCTCAAGTT 360
Db 301 gatattagatagcgttgcgaaatttcataatgataagactttgaaactctcaagtt 360
OY 361 CATTTTCTTACGACACTACCTCAATTCGTAAGATGTTTCGAAGATCCTTATATGATATAA 420
Db 361 cattttctttagcgaactacacctgaataatgctgaagaatcgtttatgcatataa 420
OY 421 ACATATTAATATGATGATCATATACCATCTGACTTCATGTTGTATGAGCCTGTTGAT 480
Db 421 acatatttaaatatgatacatatgaacccaatccgacttcaatgctgaagccttctgat 480
OY 481 GTTGTTTTATACATGAGCCCAATGTCCTGATGCGTTGCCAAAATTTGTTGTTTAAA 540
Db 481 gttgttttatacatgaccacaatgtgcctgagtgctgtcccaaatgtgttttaaa 540
OY 541 AAACGATTTGAAGCTATCCACACAATTTGATAGTACTTGAATTCACAGCATATATACGA 600
Db 541 aaacgatttgaagcattccacacaatttgaatagtacttgaataatccagaagatataaga 600
OY 601 TGGCCTTTGACAGGCGTGGCAAGCCAGCTTGTGTGGGAGCATCTCCAAAATCGGAT 660
Db 601 tggcctttgacagggctggcaagccagcttgtgtgtgggagcatctccaaaatcggat 660
OY 661 CTGCTTCCGCGTGCATCCCGCAATTCGCGGCTGACCTGAGCGCGCGCGCTGA 713
Db 661 ctgcttccgcggtgcattcccgcaatttcggttccggtccgagacgcga 713

```

RESULT 4
 Z60645
 ID Z60645 standard; DNA: 3591 BP.
 XX
 AC Z60645;
 XX
 DT 16-May-2000 (first entry)
 XX
 DE DNA encoding a fusion of aminopolylol amine oxidase/fumonisin esterase.
 XX
 KW Aminopolylol amine oxidase; APO; mycotoxin degradation; antimicrobial;
 KW plant fungal invasion; hydrogen peroxide; Fusarium; fungi;
 KW fumonisin esterase; ss.
 XX
 OS Synthetic.
 OS Exophiala spinifera.
 XX
 FH Key
 FT CDS Location/Qualifiers
 I..3691

```

FT FT /*tag= a
FT FT /product= "GST/aminopolylol amine oxidase/fumonisin
FT FT esterase fusion"
FT FT misc-feature
FT FT 1..687
FT FT /*tag= b
FT FT /note= "encodes GST and a linker"
FT FT mat_peptide
FT FT 688..2163
FT FT /*tag= c
FT FT /product= "bacterial fumonisin esterase"
FT FT 2164..2199
FT FT /*tag= d
FT FT /note= "spacer sequence"
FT FT mat_peptide
FT FT 2200..3588
FT FT /*tag= e
FT FT /product= "aminopolylol amine oxidase"
FT FT /note= "the protein has an amino terminal Lys
FT FT for optimized expression"

```

```

PD PD WO200004160-A1.
PD PD 27-JAN-2000.
PD PD 08-JUL-1999; 99WO-US15455.
PR PR 15-JUL-1998; 98US-0092936.
PA PA (PION-) PIONEER HI-BRED INT INC.
PA PA (CURA-) CURAGEN CORP.

```

```

PI PI Crasta OR, Duvick J, Folkerts O, Gilliam JT, Maddox JR;
XX XX WPI; 2000-182426/16.
DR DR P-PSDB; Y68853.
XX XX

```

```

PT PT New nucleic acid encoding aminopolylol amine oxidase, used, e.g. to
PT generate plants resistant to Fusarium

```

Example 13; Page 119-123; 145pp; English.

The present sequence encodes a fusion protein of GST/Exophiala spinifera aminopolylol amine oxidase (APO)/Fumonisin esterase. The APO enzyme has homology to the flavin containing amine oxidase family, that oxidize primary amine to an aldehyde or ketone, releasing ammonia and hydrogen peroxide. The APO enzyme degrades mycotoxins CC that promote fungal invasion of plants. Destruction of mycotoxins by APO generates, as a by-product, hydrogen peroxide which is itself an antimicrobial and stimulates the plants own defensive systems. The APO polynucleotides are used to generate plants (particularly maize) that are resistant to Fusarium or other fungi that produce mycotoxins and/or to degrade such mycotoxins (e.g. during ensilaging); for recombinant production of APO polypeptides; as selection markers for plant transformation; and to isolate related sequences from other organisms. The APO polypeptides are used to degrade mycotoxins in plant materials, including expression in engineered bacteria and fungi, e.g. rumen microflora.

Sequence 3591 BP; 758 A; 1017 C; 1064 G; 752 T; 0 other;

Query Match 62.4%; Score 684.2; DB 21; Length 3591;
 Best Local Similarity 97.5%; Pred. No. 1.4e-178;
 Matches 695; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

```

OY 1 ATGTCCCTATACAGTATTGAAATTAAGGCGCTTGTCACACCCAGCTGCTT 60
Db 1 atgtccctatactagttatgtgaataatgaaggcctgtgcacccctgactctt 60
OY 61 TTGGAATATCTTGAGAAATAATGAGAGCATTTGTATGACGGATGAGAGTATAA 120
Db 61 ttggaatattcttgaagaataaataagaagcatttgtatgagcgcgaatgaatataa 120
OY 121 TCGCAAAACAAAAGTTTGAATTTGGTTTGGAGTTTCCCAATCTTCCTATTATTGAT 180

```

```

Db 121 tggcgaaacaaagtctgaattggtgttgaggttcccaatcttcctattatattgat 180
Qy 181 GGTGATGTTAAATTAACACAGTCTATGCGCATCATGATATATAGCTGACAGCAAC 240
  |||
Db 181 ggtgatgttaataaataacacagctctatgacatcatatagctgacagcaaac 240
Qy 241 ATGTTGGGTGTTGTTCCAAAAGAGCTGACAGATTTCAATGCTTGAAGAGCGGTTTGG 300
  |||
Db 241 atgttggtgtgtgtctcaaaagagcgtgcagagattcaatgcttgtaagagcggttttg 300
Qy 301 GATATTAGTAGCGGTGTTTCGAGAAATGATATAGTAAGACTTTGAACCTCCAAAGTT 360
  |||
Db 301 gatattagatacgggtgtctcgaaatcatagataagaacttgaaacttccaaagtt 360
Qy 361 GATTTTCTTAGCAAGCTACCTGAAATGCTGAAATGTTGGAAGANTGTTTATGATATAA 420
  |||
Db 361 gatttctctagcaagctacctaagaatgctgaaatgttcgaagatcgtttatgataaa 420
Qy 421 ACATATTTAAATGCGATCATGTATACCCATCCGACTTCAATGTTATGACGCTTTGAT 480
  |||
Db 421 acatattaaatggtatcatgatgtaaccatccctgacttcatgtgtatgacgtcttgat 480
Qy 481 GTTGTATTACATGAGCAACCAATGTCGATGCGTCCCAAAATTAAGTTTGTATAA 540
  |||
Db 481 gtgttttataacatggaacccaatgctgagctgctcccaaaatagttgttataaa 540
Qy 541 AAACGATTGAGCTATGCCACAAATTAATGATTAAGTACTTGAATCCAGCAAGTATATAGA 600
  |||
Db 541 aaacgattggaagctatcccaacaattgataagctacttgaaatccagcaagatataagca 600
Qy 601 TGGCGTTTTCAGGCGTGGCAAGCCAGCTTTGTTGTCGCGACCAATCTCCAAAATCGAAT 660
  |||
Db 601 tggcgttttcagggcgtgcaagccagctgttgggtggtgacacatccctccaaatcgat 660
Qy 661 CTGGTCCGCGTGAGATCCCGAATTCGCGGTGACTCGACGCGCGCGTGA 713
  |||
Db 661 ctggttcggtgagatcccggaattccacgagatttcggtccgagagccga 713

```

RESULT 5

Z32916
Z32916 standard; cDNA: 2475 BP.

AC Z32916;

DT 09-FEB-2000 (first entry)

DE Human HCP/GST fusion gene.

KW Haematopoietic cell phosphatase; HCP; cloning; expression;

KW epithelial cell; growth factor receptor; interleukin-3; IL-3;

KW erythropoietin; EPO; negative regulator; signal transduction;

KW modulation; activity; inhibitor; stimulation; cytokine therapy;

KW antisense therapy; gene therapy; treatment; diagnosis;

KW glutathione-S-transferase; fusion; ds.

OS Synthetic.

XX Homo sapiens.

PH Key location/Qualifiers

FT CDS 1..2475 /tag= a

FT /product= "Human HCP/GST fusion protein"

XX MO9954450-A1.

XX 28-OCT-1999.

XX 15-APR-1999; 99WO-US08228.

XX 17-APR-1998; 98US-0082072.

XX (ORTH) ORTHO-MCNEIL PHARM INC.

```

XX Jolliffe L, Barbone F, Shanahan M, Xu D;
PI MPI: 2000-013244/01.
DR P-PSDB; Y52288.
XX Assay methods for modulators of haematopoietic cell phosphatase -
PT Example 3; Fig 4A; 57Pp; English.
PS
XX This sequence represents a full-length human haematopoietic
CC cell phosphatase (HCP)/glutathione-S-transferase (GST)
CC fusion gene. This was cloned and expressed in E. coli and
CC subsequently purified for use in a variety of activity assays.
CC HCP is a 68 kD non-membrane associated protein found primarily in
CC haematopoietic cells, and to some extent in epithelial cells. It
CC has been shown to associate with several growth factor receptors
CC following ligand stimulation and acts as a negative regulator
CC of receptor activation, and hence of haematopoietic signal transduction.
CC The receptors it regulates include the interleukin-3 (IL-3) and
CC erythropoietin (EPO) receptors. The invention relates to a novel
CC method of identifying compounds that modulate HCP activity. This
CC comprises combining a modulator of HCP protein activity with HCP protein
CC and HCP protein substrate, and measuring an effect of the modulator on
CC the activity of the HCP protein. Inhibitors of HCP can be used to
CC modulate the activity of haematopoietic receptors. HCP has been shown to
CC be a negative regulator of one or more signal transduction pathways in
CC haematopoietic cells, and the identification of a HCP inhibitor would
CC provide a synthetic stimulator to the haematopoietic system could be
CC used in conjunction with other cytokine therapy, e.g., administration of
CC EPO. HCP is also associated with several growth factors found in
CC haematopoietic cells and it is possible that a HCP inhibitor may also
CC function as a haematopoietic cell potentiator, enhancing the effect of
CC growth factors, decreasing the necessary dose of the factor. The HCP
CC polynucleotides are sources of probes and primers, and can be used to
CC design antisense sequences, and in gene therapy. The protein, or
CC active fragments thereof, may be used to treat or diagnose disorders
CC which require the modulation of the HCP activity. Small molecules that
CC inhibit the action of HCP can also augment the effect of EPO, which
CC meaning that a lower therapeutic dose of EPO may be utilised.
CC
XX
SQ Sequence 2475 BP; 638 A; 612 C; 701 G; 524 T; 0 other;

```

Query Match 62.4%; Score 684; DB 21; Length 2475;
Best Local Similarity 99.3%; Pred. No. 1.4e-178;
Matches 687; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

Qy 1 ATGTCCCTATAGTATATGGAATTAAGGCGCTTGTGCAACCACTGCACTTCTT 60
  |||
Db 1 atgtccctatactagtagtataaataaagagccttgtaacccactgcactctt 60
Qy 61 TTGGAATATCTTGAAGAAAATATGAAGAGCATTTGTATGAGCGCGATGAGGTGATATA 120
  |||
Db 61 ttggaatattctgaagaaaatatagaagcatttgatagagcgatgaagtgataaa 120
Qy 121 TGGCGAACAACAAAAGTTGATGAGGTTGGAGTTTCCCAATCTTCTTATTTATGAT 180
  |||
Db 121 tggcgaaacaaagtttgatgaggttggtgagtttcccaatcttccctattatattgat 180
Qy 181 GGTGATGTTAAATTAACACAGTCTATGCGCATCATGATATATAGCTGACAGCAAC 240
  |||
Db 181 ggtgatgttaataaataacacagctctatgacatcatatagctgacagcaaac 240
Qy 241 ATGTTGGGTGTTGTTCCAAAAGAGCTGACAGATTTCAATGCTTGAAGAGCGGTTTGG 300
  |||
Db 241 atgttggtgtgtgtctcaaaagagcgtgcagagattcaatgcttgtaagagcggtttg 300
Qy 301 GATATTAGTAGCGGTGTTTCGAGAAATGATATAGTAAGACTTTGAACCTCCAAAGTT 360
  |||
Db 301 gatattagatacgggtgtctcgaaatcatagataagaacttgaaacttccaaagtt 360
Qy 361 GATTTTCTTAGCAAGCTACCTGAAATGCTGAAATGTTGGAAGANTGTTTATGATATAA 420

```

```

Db      361 gatttcttagcaagcaccggaatgctgaaatgctgagaatgcttattgcaataa 420
Oy      421 ACATATTAAATGSGTATCATGTAACCCATCCGACTTCATGTTGATGACGCTTGAT 480
Db      421 acataatlaaattgctgctatgtaacccatccctactctgtglatgagcgtctgatt 480
Oy      481 GTTGTTTATACATGAGCCCATGTGCTGATGCTTCCCAAAATAGTTGTTTAA 540
Db      481 gtttcttatacaaggacccaatgctgctgagcgtctcccaaatagtttttaa 540
Oy      541 AACGTTTAAAGTATCCCAAAATGATAGTACTGTAATTCAGCAAGTATATAGCA 600
Db      541 aaacgatttgaagctatcccaaatgataagtaacttgaaatccagcaagatataagca 600
Oy      601 TGGCCTTTGACAGGCTGACAGCCAGCTTGTGTGGGCGACCATCTCCAAATCGGAT 660
Db      601 tggcctttgagagcgtgagcaagcaagcttgggtggtggcgacatccccaatccgaat 660
Oy      661 CTGGTTCCGCTGATCCCGAATTCCTCCGGGT 692
Db      661 ctggtccgctgtgattcccggaattcatggt 692

```

RESULT 6

Z32917 232917 standard; cDNA; 1866 BP.

XX Z32917;

DT 09-FEB-2000 (first entry)

DE Human HCP phosphodomain/GST fusion gene.

XX Haematopoietic cell phosphatase; HCP; cloning; expression;

KW epithelial cell; growth factor receptor; interleukin-3; IL-3;

KW erythropoietin; EPO; negative regulator; signal transduction;

KW modulation; activity; inhibitor; stimulation; cytokine therapy;

KW antisense therapy; gene therapy; treatment; diagnosis; phosphodomain;

KW catalytic domain; glutathione-S-transferase; fusion; ds.

XX Synthetic.

OS Homo sapiens.

XX Key

PF 15-APR-1999; 99WO-US08228.

PR 17-APR-1998; 98US-0082072.

XX (ORTH) ORTHO-MCNEIL PHARM INC.

PI Jolliffe L, Barbone F, Shanahan M, Xu D;

DR MPI; 2000-013244/01.

XX P-PSDB; Y52289.

XX Assay methods for modulators of hematopoietic cell phosphatase

XX Example 2; Fig 4C; 57pp; English.

CC This sequence represents a human haematopoietic cell

CC phosphatase (HCP) catalytic domain/glutathione-S-transferase (GST)

CC fusion gene. This was cloned and expressed in E. coli and

CC subsequently purified for use in a variety of activity assays.

CC HCP is a 68 kD non-membrane associated protein found primarily in

```

CC haematopoietic cells, and to some extent in epithelial cells. It
CC has been shown to associate with several growth factor receptors
CC following ligand stimulation and acts as a negative regulator
CC of receptor activation, and hence of haematopoietic signal transduction.
CC The receptors it regulates include the interleukin-3 (IL-3) and
CC erythropoietin (EPO) receptors. The invention relates to a novel
CC method of identifying compounds that modulate HCP activity. This
CC comprises combining a modulator of HCP protein activity with HCP protein
CC and HCP protein substrate, and measuring an effect of the modulator on
CC the activity of the HCP protein. Inhibitors of HCP can be used to
CC modulate the activity of haematopoietic receptors. HCP has been shown to
CC be a negative regulator of one or more signal transduction pathways in
CC haematopoietic cells, and the identification of a HCP inhibitor would
CC provide a synthetic stimulator to the haematopoietic system could be
CC used in conjunction with other cytokine therapy, e.g., administration of
CC EPO. HCP is also associated with several growth factors found in
CC haematopoietic cells and it is possible that a HCP inhibitor may also
CC function as a haematopoietic cell potentiator, enhancing the effect of
CC growth factors, decreasing the necessary dose of the factor. The HCP
CC polynucleotides are sources of probes and primers, and can be used to
CC design antisense sequences, and in gene therapy. The protein, or
CC active fragments thereof, may be used to treat or diagnose disorders
CC which require the modulation of the HCP activity. Small molecules that
CC inhibit the action of HCP can also augment the effect of EPO, which
CC meaning that a lower therapeutic dose of EPO may be utilised.
CC
XX
SO

```

Sequence 1866 BP; 515 A; 444 C; 503 G; 404 T; 0 other;

Query Match 62.4%; Score 683.4; DB 21; Length 1866;
Best Local Similarity 99.1%; Pred. No. 1.8e-178;
Matches 687; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

Oy      1 ATGTCCCTATACATGCTTATGGAATTAAGGCGCTTGACACCCACTGACTTCTT 60
Db      1 atgtccctatcactagcttattggaatlaaaggccttggacaccacactgctctt 60
Oy      61 TTGCAATATCTTGAAGAAATATGAAGAGCATTTGTATGCGCGGATGAAAGTGAT 120
Db      61 ttgcaatctcttgaagaaaataagaagacattgtatgaagcgatgaagtgataaa 120
Oy      121 TGGCGAAACAAAGATTGAAATGGGTTTGAGATTGCCAATCTCCCTATTTATTTAT 180
Db      121 tggcgaaacaaaagattgaaattgggttggatttgcacattctccattcttatatgat 180
Oy      181 GGTGATGTTAAATTAACAGCTCATGCGCATCATATAGCTTATAGTGCAAGCACAC 240
Db      181 ggtgattgtaatttaacacagctcatgcatcatagctatagctgacagcaaac 240
Oy      241 ATGTGGGTGGTGTGCCAAAAGCGTGACAGATTTCAATGCTTGAAGAGGGGTTTG 300
Db      241 atgtgggtgtgtgtgccaaaagcgtgacagatttcaatgcttgaagagggttttg 300
Oy      301 GATATTTAGATACGCTGTTTGAGCAATTTGCAATPATAAGACTTTGAAACTCAAG 360
Db      301 gatattagatcagctgttttgagcaatttgcaatpataagactttgaaactccaagt 360
Oy      361 GATTTCTTAGCAAGCTACCTGGAATGCTGAAATGTTGCAAGATGTTATATTCATA 420
Db      361 gatttcttagcaagcctacctggaatgctgaaatgctgaaatgcttattgctataaa 420
Oy      421 ACATATTAAATGSGTATCATGTAACCCATCCGACTTCATGTTGATGACGCTTGAT 480
Db      421 acataatlaaattgctgctatgtaacccatccctactctgtglatgagcgtctgatt 480
Oy      481 GTTGTTTATACATGAGCCCATGTGCTGATGCTTCCCAAAATAGTTGTTTAA 540
Db      481 gtttcttatacaaggacccaatgctgctgagcgtctcccaaatagtttttaa 540
Oy      541 AACGTTTAAAGTATCCCAAAATGATAGTACTGTAATTCAGCAAGTATATAGCA 600
Db      541 aaacgatttgaagctatcccaaatgataagtaacttgaaatccagcaagatataagca 600

```

```
OY 601 TGGCCTTTGACGGCTGGACACGCTTGGTGGTGGGACACCTCCAAATGCGAT 660
    |||||||
DB 601 Tggccttgcagagcgcgcgaacccagcttggcgcgcgcacatcccaaatcgat 660
OY 661 CTGGTTCGCGGTGATCCCGGAATTCGCGGTC 693
    |||||||
DB 661 ctggttcgcgcgtgacatcccggaattcgcgcc 693

RESULT 7
ID 258400 standard; cDNA; 2079 BP.
XX 258400;
AC
XX
XX
XX 23-MAY-2000 (first entry)
DE DNA encoding GST-truncated amino polyol amine oxidase fusion.
XX
XX Amino polyol amine oxidase; APAO; fumonisin; mycotoxin;
KW transgenic plant; detoxification; animal feed; silage;
KW selectable marker; glutathione S transferase; GST; ss.
XX
OS Chimeric - Schistosoma japonicum.
OS Chimeric - Exophiala spinifera.
XX
XX Key Location/Qualifiers
FH sig_peptide 1..687
FT /tag= a
FT /product= "gst fusion + polylinker"
FT mat_peptide 688..2076
FT /tag= b
FT /product= "K:trPAO"
FT misc_feature 688..690
FT /*tag= c
FT /*note= "extra lysine"
XX
XX WO200004159-A1.
XX
XX 27-JAN-2000.
XX
XX 08-JUL-1999; 99WO-US15454.
XX
XX 15-JUL-1998; 98US-0092936.
XX
XX 21-MAY-1999; 99US-0135391.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX (CURA-) CURAGEN CORP.
XX
XX Duvalck JP, Gilliam JT, Maddox JR;
XX
XX WPI, 2000-182425/16.
XX
XX P-PSDB; Y58911.
XX
XX New isolated polynucleotides, polypeptides useful for detecting and
PT degrading fumonisin or structurally related mycotoxin in processed
PT grain or in silage -
XX
XX Example 9; Page 89-92; 154pp; English.
XX
XX The present sequence is that of a polynucleotide encoding a
CC GST-K:trPAO fusion protein composed of glutathione S transferase
CC and K:trPAO, a truncated, but functional, amino polyol amine
CC oxidase (see also Y58909) of Exophiala spinifera. The construct
CC was designed for expression of K:trPAO in Escherichia coli.
CC K:trPAO is capable of degrading fumonisin and related mycotoxins.
CC The invention provides APAO polynucleotides (see Z58383-87) and
CC polypeptides (see Y58900-05) of E. spinifera and Rhinocladiella
CC atrovirens. The polynucleotides are used to transform plant cells
CC normally susceptible to Fusarium or other toxin-producing fungus
CC infection. Also provided are methods for expressing APAO in
CC transgenic plants, prokaryotic and non-plant eukaryotic systems.
CC Methods for detoxification of grain, grain processing, silage, food
```

```
CC crops and in animal feed and rumen microorganisms are also disclosed.
CC APAO polynucleotide can also be used as a selectable marker.
XX
XX Sequence 2079 BP; 537 A; 490 C; 547 G; 505 T; 0 other;
```

```
Query Match 62.3%; Score 683.2; DB 21; Length 2079;
Best Local Similarity 98.9%; Pred. No. 2.1e-178;
Matches 688; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

```
OY 1 ATGTCCCTATCTAGCTATTTGAAATTTAAGGCGCTGTGCAACCCACTGACTTCTT 60
    |||||||
DB 1 atgtccctatactagctatttggaaatttaaggcgctgtgcaacccactgacttctt 60
OY 61 TTGAATATCTTGAACAAAATTTGAGAGCATTTGTATGAGCGCGATGAGTGATAA 120
    |||||||
DB 61 ttggaatattcttgaacaaaatttgagagcatttgtatgagcgcgatgagtgataa 120
OY 121 TGGCGAACAACAAAGTTTGAAATTTGGATTTGCCAATCTTCTATTATATTTGAT 180
    |||||||
DB 121 tggcgaaacaacaaagtlttgatgattggattggatttcccaatcttctattatattgt 180
OY 181 GGTGATGTTAAATTTAACACAGTCTATGCCATCATACGTTATATACCTGACAACAC 240
    |||||||
DB 181 ggtgatgtttaaatttaacacagctctatgcatcatatagcttgaacaagcacaac 240
OY 241 ATGTTGGTGGTGTGTTGCCAAGAGCGTGCAGAGATTTCATGCTTGATGAGAGCGTTTG 300
    |||||||
DB 241 atgttggtggtgtgttgcgaagagcggtgcagagatttcgatgagagcggttttg 300
OY 301 GATATTAGATACGCTGTTTCGAGAAATTTGATATAGTAAGACTTTGAAACTCTCAAGATT 360
    |||||||
DB 301 gatattagatacggtgttccgaaatgttgatatagtaaagactttgaaactccaagatt 360
OY 361 GATTTTCTTAGCAAGCTACTGAAATGCTGAAATTTGAGAGATGTTATGTCTATAA 420
    |||||||
DB 361 gattttcttagcaagctactgaaatgctgaaatgttgagagatgttatgtctataaa 420
OY 421 ACATATTTAAATGATGATCATGTAACCCATCCGACTTCATGTTATGAGCGCTTGAT 480
    |||||||
DB 421 acatatttaaattgattcatgttaacccatccgacttcattgattgagcgcttgat 480
OY 481 GTTGTTTTATACATGACCAACCAATGTGCTGATGCTGTCCTCCAAATATGTTGTTTAA 540
    |||||||
DB 481 gttgttttatacatgacccaaccaatgtgctgatgctgctcccaaatatgttgttttaa 540
OY 541 AAACGATTGAAAGCTATCCCAACAATTGATAGTACTTGAATCCAGCAAGTATATAGA 600
    |||||||
DB 541 aaacgattgaaagctatcccaacaattgatagtacttgaatccagcaagtatatagca 600
OY 601 TGGCCTTTGACGGCTGGACACGCTTGGTGGTGGGACACCTCCAAATGCGAT 660
    |||||||
DB 601 Tggccttgcagagcgcgcgaacccagcttggcgcgcgcacatcccaaatcgat 660
OY 661 CTGGTTCGCGGTGATCCCGGAATTCGCGGTCGAC 696
    |||||||
DB 661 ctggttcgcgcgtgacatcccggaattcgcgcc 696

RESULT 8
ID 260639 standard; DNA; 2079 BP.
XX 260639;
AC
XX
XX 16-MAY-2000 (first entry)
DE DNA encoding an aminopolyol amine oxidase/GST fusion.
XX
XX Aminopolyol amine oxidase; APAO; mycotoxin degradation; antimicrobial;
KW plant fungal invasion; hydrogen peroxide; Fusarium; fungi; ss.
XX
XX Synthetic.
```

Dd	181	ggtgacgtttaaataaacacagtcctatggccatcatcagcttatatagctgcagcaagcacac	240
Oy	241	ATGTTGGGTGGTGGTCACAAAGAGGCTGCAGAGATTCAATGCTTGAAAGAGCGGTTTTG	300
Dd	241	atgttgvgtggltgtgccaaaagaagcgvtgcagagattccaigtctgaaggagcggtlltg	300
Oy	301	GATATTAGATACGGTGTTCGAGAATTCATATAGTAAGAATTGAAGACTTCAGAAATT	360
Dd	301	gatattagaatacgttgttcggaattgcatalagttaagaactttgaaacctccaagtc	360
Oy	361	GATTTCTTAGAAGACTACCTCCGTAATGCTGGAANAATGTCGAAGTCGTTATGTCATAAA	420
Dd	361	gatttctctaagaacgtaccgtaaatgctgaanaatglttcgaagatcgcttatgtcataaa	420
Oy	421	ACATAATTAATAGTGATCATGTATAACCCATCCTGACTTCATGTTGTATGAGCGCTTGAT	480
Dd	421	acataattaacgtgtgcatatglaaacacatcctgactcatgtgtatgtagcgtctgat	480
Oy	481	GTMTGTTTATACATGAGACCACATGTCCTGSATGCGTTCCCAAATTAGTTGTTTTAAA	540
Dd	481	gttgttctataacatgagaccacatgycctgtgagtggttccccaaatatagtlgtltaa	540
Oy	541	AAACGATTGAAGCTATCCCACAATTATATAGTAACTTGAAATCCACAGTATATAGCA	600
Dd	541	aaacgatttgaagcgtatcccacaatatgtatagtaacttgaatccagaagtataagca	600
Oy	601	TGGCCTTTGACAGSGCTGCCAAGCACGTTTGTGTGTGCGCACATCTCCMAAATCGAT	660
Dd	601	tggcctttgcagaggtgcgtgcaagcaacglttgtgtggtgcacatcctccaatatcgat	660
Oy	661	CYGTTCGCGCGTGAGTCCCGCAATTCGCGGTGCAC	696
Dd	661	ctggttcggtgtgatatcccggaattcaaagaacac	696
RESULT	9		
ID	Z58407	standard; cDNA; 2490 BP.	
XX	Z58407;		
AC			
XX			
Dt	23-MAY-2000	(first entry)	
XX			
DE	DNA encoding GST-truncated amino polyol amine oxidase fusion.		
XX			
KW	Amino polyol amine oxidase; APAO; funonisin; myco toxin;		
KM	transgenic plant; detoxification; animal feed; silage;		
KW	selectable marker; glutathione S transferase; GST; mutant; ss.		
XX			
OS	Chimeric - Schistosoma japonicum.		
OS	Chimeric - Exophiala spinifera.		
OS	Synthetic.		
XX			
FH	Key	Location/Qualifiers	
FT	sig_peptide	1..687	
FT		/*tag= a	
FT		/product= "gst fusion + polylinker"	
FT	mat_peptide	688..2076	
FT		/*tag= b	
FT		/product= "K::trAPAO"	
FT	misc_feature	688..690	
FT		/*tag= C	
FT		/note= "extra lysine"	
FT	mutation	replace(1288..1290, TCC)	
FT		/*tag= d	
FT	mutation	replace(1303..1305, AAC)	
FT		/*tag= e	
XX			
PN	MO200004159-A1.		
PD	27-JAN-2000.		
XX			
FE	08-JUL-1999.	99WO-US15454.	


```

XX 15-JUL-1998; 98US-0092936.
PR 21-MAY-1999; 99US-0135391.
XX (PION-) PIONEER HI-BRED INT INC.
PA (CURA-) CURAGEN CORP.
XX DuVick JP, Gilliam JT, Maddox JR;
XX WPI: 2000-182425/16.
DR P-PSDB: Y58918.
XX
XX New isolated polynucleotides, polypeptides useful for detecting and
PT degrading fumonisin or structurally related mycotoxin in processed
PT grain or in silage
XX
XX Example 15; Page 131-135; 154pp; English.
XX
XX The present sequence is that of a polynucleotide encoding a
CC GST:APAO fusion protein (see Y58918) composed of glutathione S
CC transferase and Exophiala spinifera amino polyol amine oxidase
CC (APAO). 2 Codons of the APAO coding region have been altered by
CC site-directed mutagenesis in order to change the glycosylation
CC pattern of the fusion protein. These were the replacement of AAT
CC (Asn) by TCC (Ser) at codon 430, and replacement of AGC (Ser) by
CC AAC (Asn) at codon 435. APAO is capable of degrading fumonisin and
CC related mycotoxins. The invention provides APAO polynucleotides
CC (see Z58383-87) and polypeptides (see Y58900-05) of E. spinifera
CC and Rhinocladella atrovirens. The polynucleotides are used to
CC transform plant cells normally susceptible to Fusarium or other
CC toxin-producing fungus infection. Also provided are methods for
CC expressing APAO in transgenic plants, prokaryotic and non-plant
CC eukaryotic systems. Methods for detoxification of grain, grain
CC processing, silage, food crops and in animal feed and rumen
CC microorganisms are also disclosed.
XX
XX Sequence 2490 BP; 625 A; 623 C; 653 G; 589 T; 0 other:
XX
XX Query Match 62.3%; Score 683; DB 21; Length 2490;
XX Best Local Similarity 99.3%; Pred. No. 2.6e-178;
XX Matches 686; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
XX
XX 1 ATGTCCTATATAGTATTGAAATTAAGGCGCTTGCAACCCAGCTGCTTCTT 60
XX 1 atgtccctatactagtatttgaaataaaggccttgcaaccactgcactctct 60
XX 61 TTGGAATATCTTGAGAAAATATAGAGCATTTGTATGACCGGATGAAGTATAA 120
XX 61 ttggaatatcttgagaaaataatagagcatttgtagagcgagatgaagtgataaa 120
XX 121 TGGCAAAACAAAAGTTGAAATGGATTGGATTTCCCAATCTTCCTATTATTTGAT 180
XX 121 tggcaaaaacaaaagttgaaattggattggagtttcccaatcttccctattatattg 180
XX 122 tggcgaaaacaaaagttgaaattggattggagtttcccaatcttccctattatattg 180
XX 181 GGTGATGTTAAATTAACACAGTCTATGGCCATCATAGCTTATATAGCTGACACAAAC 240
XX 181 ggtgatgtttaaattaacacagctctatggccatcatagcttataatgctgaagcacaac 240
XX 181 ggtgatgtttaaattaacacagctctatggccatcatagcttataatgctgaagcacaac 240
XX 241 ATGTTGGGTGGTGTTCGAAAAGAGCGTGAGATTTCAATGCTTGAAGGACGGTTTG 300
XX 241 atgttgggtggtgttcgaaaagagcgctgagatttcaatgcttgaaggacggtttg 300
XX 241 atgttgggtggtgttcgaaaagagcgctgagatttcaatgcttgaaggacggtttg 300
XX 301 CATATTAGATACGGTGTTCGAGAAATGATATAGTAAAGACTTTGAAACCTGCAAGTT 360
XX 301 catattagatacgggtgttcgagaaatgatataagtaaaagacttgaacacttcaagtt 360
XX 301 gatattagatagcggtgttcgagaaatgatataagtaaaagacttgaacacttcaagtt 360
XX 361 GATTTTCTTAGACACTACCTGAATGCTGAATAATGTTGAAGATCGTTTATGTCATAAA 420
XX 361 gattttcttagacactacctgaatgctgaataatgttgaagatcgtttattgctataaa 420
XX 361 gattttcttagacactacctgaatgctgaataatgttgagatcgtttattgctataaa 420
XX 421 ACATATTTAAATGATGATCATGTATACCCATCTGACTTCATTTGATATACGCTCTTGT 480
XX 421 acatattttaaataatgatacatgtataccctctgacttcatttggatatacgccttgt 480

```

```

DB 421 acatatttaaattgtagatcatgtaaccacactgacttcatgttgaagcgtcttgat 480
OY 481 GTTGTATTATACATGAGCCCAATGTCGTGATGCTTCCCAAAATTAAGTTGTTTAA 540
DB 481 gtgtgtttatatactgagcccaatgctgctgagatgcttcccaaatatgattttaa 540
OY 541 AAACGATTTGAAGCTATCCACAAATTAAGTACTTGAATAATCCAGCAATATATAGCA 600
DB 541 aaacgattggaagctatcccaaatgataagtaacttgaatccagcaagtatatagca 600
OY 601 TGGCCTTTGACAGGCTGCGAAGCCAGCTTTGTGTGGCGACCATCTCCAAATGGAT 660
DB 601 tggcctttgacaggctgcggaagccagcttgtgtggcgaccatctccaaatcgat 660
OY 661 CTGCTTCCGCGTGATGCCCGAATTCGCGG 691
DB 661 ctgcttccgcgtgatacccggaattcatcg 691

RESULT 10
ID 260646
XX 260646 standard; DNA: 2490 BP.
XX
XX 260646:
XX
XX 16-MAY-2000 (first entry)
XX
XX DNA encoding a fusion of GST/mutated aminopolylol amine oxidase.
XX
XX Aminopolylol amine oxidase; APAO; mycotoxin degradation; antimicrobial;
XX plant fungal invasion; hydrogen peroxide; Fusarium; fungi; ss.
XX
XX Synthetic.
XX
XX Exophiala spinifera.
XX
XX Key Location/Qualifiers
XX CDS 1..2490
XX FT /tag= a
XX FT /product= "GST/aminopolylol amine oxidase fusion"
XX FT misc-feature 1..687
XX FT /tag= b
XX FT /note= "encodes GST and a linker"
XX FT mat-peptide 688..2487
XX FT /tag= c
XX FT /product= "aminopolylol amine oxidase"
XX FT /note= "the protein has been mutated to remove
XX potential glycosylation sites at positions
XX 201 and 206"
XX
XX WO200004160-A1.
XX
XX 27-JAN-2000.
XX
XX 08-JUL-1999; 99WO-US15455.
XX
XX 15-JUL-1998; 98US-0092936.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX (CURA-) CURAGEN CORP.
XX
XX Crasta OR, DuVick J, Folkerts O, Gilliam JT, Maddox JR;
XX WPI: 2000-182426/16.
XX DR P-PSDB: Y68854.
XX
XX New nucleic acid encoding aminopolylol amine oxidase, used, e.g. to
XX generate plants resistant to Fusarium
XX
XX Example 15; Page 127-130; 145pp; English.
XX
XX The present sequence encodes a fusion protein of GST and a mutated
XX Exophiala spinifera aminopolylol amine oxidase (APAO). The APAO
XX protein has been mutated to remove potential glycosylation sites at

```

CC position 201 and 206. The APO enzyme has homology to the flavin
CC containing amine oxidase family, that oxidise primary amine to an
CC aldehyde or ketone, releasing ammonia and hydrogen peroxide. The APO
CC enzyme degrades mycotoxins that promote fungal invasion of plants.
CC Destruction of mycotoxins by APO genes, as a by-product, hydrogen
CC peroxide which is itself an antimicrobial and stimulates the plants
CC own defensive systems. The APO polynucleotides are used to generate
CC plants (particularly maize) that are resistant to Fusarium or other
CC fungi that produce mycotoxins and/or to degrade such mycotoxins
CC (e.g. during ensiling); for recombinant production of APO
CC polypeptides, as selection markers for plant transformation; and to
CC isolate related sequences from other organisms. The APO polypeptides
CC are used to degrade mycotoxins in plant materials, including expression
CC in engineered bacteria and fungi, e.g. rumen microflora.
XX
Sequence 2490 BP; 635 A; 623 C; 653 G; 589 T; 0 other;

Query Match	62.3%	Score 683	DB 21	Length 2490
Best Local Similarity	99.3%	Pred. NO. 2.6e-178		
Matches 686	Conservative 0	Mismatches 5	Indels 0	Gaps 0
QY 1	ATGTCCTCTACTAGTGGTATTGGAAAATTGAAGGCGCTTGTCGAACCCACTGACTTCTT	60		
Db 1	atgcccctactactggtattgtggaatttaaggcccttgfcgaaccactcgaacttcct	60		
QY 61	TTGCAATATCTTGAGAGAAAATATGAGAGCATTTGTATGAGCGGATGACGTATATA	120		
Db 61	ttggaatctcttgagaagaaaatatgaagacgtcttgatgaagcgtatgaagtgataaa	120		
QY 121	TGGCGCAACAAAAGATTGGAAATGGGCTTGGAGTTTCCCAATCTTCCTATATATGAT	180		
Db 121	tggcgaacaaaaaatttggaaatggggttgcaggttcccaactcttcctattatattgt	180		
QY 181	GGTGATGTTAAATTTAAACAGCTCTATGGCCATCATACGTTATATAGCTGACAGACAAC	240		
Db 181	ggtagtgaataatcacacagctctatggccatcatcaactatataatgctgacagaacaaac	240		
QY 241	ATGTGGGTGGTGTGTCACAAAAGCGGTGCAGAGATTTCATGCTTTGAAGAGCGGTTTG	300		
Db 241	atgtgggtgggtgtgtcacaaaagcgtgcagagatttcacatgcttgaagagcgttttg	300		
QY 301	GATATTAATACATACGGGTTTTCGAGATTGGCATATATATAAAGACTTTGAACATCTCAAGTT	360		
Db 301	gatatataagacgagtttgcgagactgcatatagtaaagacttggaaactctccaagtt	360		
QY 361	GATTTCCTTACCAAGCTACCTGAAATGCTGAAAATGTTGGAAGATCGTTATATGATATAA	420		
Db 361	gatttctctagaaagctaacctcgaaatgcttgaanaatgcttcgaagaatcgtttatgctataaa	420		
QY 421	ACATATTTAAATGGTGATCATGTATAACCATCTCTGACTTCATGTTTATATAGAGCTCTTGAT	480		
Db 421	acataatlaaaatggtacatgatacgtacacccactcgacttcaatgltatgaagcgtcttgat	480		
QY 481	GTGTGTTTATACATGAGACCCAAATGCGCTGATGCGTTCGCCAAATTTGTTGTTTAA	540		
Db 481	gttgtttatatacatgagcccaatgycctgtagtggttcccaaatglttggltttaa	540		
QY 541	AAACGATTTGAGAGCTATCCACAAATTTATATAGTACTTGAAATCCAGCAAGTATATAGCA	600		
Db 541	aaacgatattgaaagctatctccacaataatgataagatcttgaatccagcaagtatatagca	600		
QY 601	TGGCCTTTGCAAGGCGTGSCAAGCACAAGCTTTGGTGGTGGCGACATCTCCAAAATCGGAT	660		
Db 601	tggcctttgcagagcgtgcaagccacgttltgltggtgcaccatctccaaaatcgcgat	660		
QY 661	CTGGTTCGGCGTGATCCCGCAATTCGCGG	691		
Db 661	ctggttccgagtgatcccggaatttcaag	691		

ID	2584405 standard; cDNA; 3618 BP.
XX	
AC	2584405;
XX	
DT	23-MAY-2000 (first entry)
XX	
DE	DNA encoding fumonisin esterase- <i>amino polyol amine oxidase fusion</i> .
XX	
KW	Amino polyol amine oxidase; APAO; fumonisin esterase; mycotoxin;
KW	transgenic plant; detoxification; animal feed; silage;
KW	glutathione S transferase; alpha-amylase; signal peptide; maize;
KW	ss.
XX	
OS	Chimeric - Schistosoma japonicum.
OS	Chimeric - Exophiala spinifera.
XX	
EH	Key
FT	sig_peptide
FT	Location/Qualifiers
FT	1..687
FT	/tag= a
FT	/product= "Gst - polylinker"
FT	677..3615
FT	/tag= b
FT	/product= "fumonisin esterase-K:trAPAO fusion"
FT	2227..2229
FT	/tag= c
FT	/note= "extra lysine codon"
XX	
PN	WO200004159-A1.
XX	
PD	27-JAN-2000.
XX	
PF	08-JUL-1999; 99WO-US15454.
XX	
PR	15-JUL-1998; 98US-0092936.
PR	21-MAY-1999; 99US-0135391.
XX	
PA	(PTON-) PIONEER HI-BRED INT INC.
PA	(CURA-) CURAGEN CORP.
PI	Duvick JP, Gilliam JT, Maddox JR;
XX	
DR	WPI; 2000-182425/16.
DR	P-PSDB; Y58916.
XX	
FT	New isolated polynucleotides, polypeptides useful for detecting and
FT	degrading fumonisin or structurally related mycotoxin in processed
FT	grain or in silage -
XX	
XX	
PS	Example 13; Page 115-120; 154pp; English.
XX	
CC	The present sequence is that of a polynucleotide encoding a
CC	fusion protein (see Y58916) composed of glutathione S transferase
CC	and Exophiala spinifera fumonisin esterase (ESF1) joined via a
CC	peptide linker to K:trAPAO (see Y58909), a truncated, but still
CC	functional, amino polyol amine oxidase of E. spinifera. The
CC	construct was designed for expression in Escherichia coli of a hybrid
CC	protein having ESPI and APAO activity. The invention provides APAO
CC	polynucleotides (see 258383-87) and polypeptides (see Y58900-05) of
CC	E. spinifera and Rhinocladella atrovirens. The polynucleotides
CC	are used to transform plant cells normally susceptible to Fusarium
CC	or other toxin-producing fungus infection. Also provided are
CC	methods for expressing APAO, optionally as a fusion protein with
CC	fumosin esterase, in transgenic plants, prokaryotic and non-plant
CC	eukaryotic systems. Methods for detoxification of grain, grain
CC	processing, silage, food crops and in animal feed and rumen
CC	microorganisms are also disclosed.
XX	
XX	Sequence 3618 BP; 884 A; 925 C; 957 G; 852 T; 0 other;

	Query Match	62.2%	Score 683.2;	DB 21;	Length 3618;
	Best Local Similarity	99.6%;	Pred. No. 5e-178;		
	Matches 684; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0;

```

QY 1 ATGTCCTTACTAGTATTGGAATAAAGGCGCTTGCAACCCACTGACTTCTT 60
DB 1 atgtccctactactagttatttggaataataaggccttgcaaccactgactctt 60
QY 61 TTGGAATATCTTGAGAAAATATAGAGCATTTGTATGACCGGATGAAGTATAA 120
DB 61 ttggaatattcttgagaaaatataagagcatttgtagagcgagtagaagtgataaa 120
QY 121 TGGCAAAACAAAGATTGGAATTGGATTGGATTTCCCAATCTTCTTATTATTGAT 180
DB 121 tggcaaaacaaaagatttggaattggattggatttcccaatcttcttattatttgat 180
QY 181 GGTGATGTTAAATTAACACAGTCTATGCCATCATACGTTATATAGCTGACACAAAC 240
DB 181 ggtgatgtttaaattaacacagtctatgscatcatagctatataagctgacaagacaac 240
QY 241 ATGTTGGGCGGTGTCNCAAAAGAGCGTGAGATTTCAATGCTTGAGAGCGGTTTG 300
DB 241 atgttggcggtgttcncaaaagagcgtagagatttcaatgcttgagagcggtttg 300
QY 301 GATTATTAGATCGTGTTCGAGAAATGCAATAGTAAGACTTTGAACCTGCTCAAAAGTT 360
DB 301 gatattagatacgtgttgcagaaatgcaatagtaagactttgaacctgctcaaaagt 360
QY 361 GATTCTTTAGCAGCTACCTGAAAATGCTGAAAATGTTGGAAGATCGTTATGATATAA 420
DB 361 gattctttagcagctacacctgaaaatgctgaaaatgttggaagatcgttatgataaa 420
QY 421 ACATATTTAAATGCGATCATGTATACCCATCTGACTTATGTTATAGCGCTTGTAT 480
DB 421 acatattttaaattgcgatcatgtataccctgacttattgtatagcgcttgtat 480
QY 481 GTTGTATTATACATGACCCCAATGTGCTGATGCGTTCCTCCAAATATTGTTTATAA 540
DB 481 gttgtattatacatgaccccaatgtgctgatgcttcctccaaatattgttttataa 540
QY 541 AAAGCTATTGAAGCTATCCACAAATGTAGTAACTTGAATTCACGAGATATATAGCA 600
DB 541 aaagctattgaagctatcccaaatgtagtaagtaacttgaattcacgagatataagca 600
QY 601 TGGCTTTGCAAGCGTGCAAGCAAGCTTGTGTGCGGACACCTCCCAAAATGCGAT 660
DB 601 tggctttgcaagcgtgcaagcaagcttgtgtgctggacacctcccaaaatgcgat 660
QY 661 CTGTTCCGCGTGATCCCGCAATTC 687
DB 661 ctggttcgcgctgagatcccggaatlc 687

RESULT 12
ID Z60644 standard; DNA: 3618 BP.
AC Z60644:
DT 16-MAY-2000 (first entry)
XX
DE DNA encoding a fusion of aminopolylol amine oxidase/fumonisin esterase.
KW Aminopolylol amine oxidase; APAO; mycotoxin degradation; antimicrobial;
  plant fungal invasion; hydrogen peroxide; Fusarium; fungi;
  fumonisin esterase; ss.
XX
OS Synthetic.
XX Exophiala spinifera.
FH Key Location/Qualifiers
FT CDS 1..3618
FT /tag= a
FT /product= "GST/aminopolylol amine oxidase/fumonisin
FT /misc-feature 1..687

```

```

FT /tag= b
FT /note= "encodes GST and a linker"
FT mat_peptide 688..2190
FT /tag= c
FT /product= "fumonisin esterase"
FT misc-feature 2191..2226
FT /tag= d
FT /note= "spacer sequence"
FT mat_peptide 2227..3615
FT /tag= e
FT /product= "aminopolylol amine oxidase"
FT /note= "the protein has an amino terminal Lys
  for optimized expression"

WO200004160-A1.
PD 27-JAN-2000.
XX
XX 08-JUL-1999; 99WO-US15455.
XX
XX 15-JUL-1998; 98US-0092936.
XX
PA (PION-) PIONEER HI-BRED INT INC.
PA (CURA-) CURAGEN CORP.
XX
XX Crasta OR, Duvick J, Folkerts O, Gilliam JT, Maddox JR;
XX WPI: 2000-182426/16.
XX P-PSDB: Y68852.
XX
XX New nucleic acid encoding aminopolylol amine oxidase, used, e.g. to
XX generate plants resistant to Fusarium.
XX
XX Example 13; Page 110-115; 145pp; English.
XX
XX The present sequence encodes a fusion protein of GST/Exophiala
XX spinifera aminopolylol amine oxidase (APAO)/fumonisin esterase.
XX The APAO enzyme has homology to the flavin containing amine oxidase
XX family, that oxidise primary amine to an aldehyde or ketone, releasing
XX ammonia and hydrogen peroxide. The APAO enzyme degrades mycotoxins
XX that promote fungal invasion of plants. Destruction of mycotoxins by
XX APAO generates, as a by-product, hydrogen peroxide which is itself an
XX antimicrobial and stimulates the plants own defensive systems. The
XX APAO polynucleotides are used to generate plants (particularly maize)
XX that are resistant to Fusarium or other fungi that produce mycotoxins
XX and/or to degrade such mycotoxins (e.g. during ensilaging); for
XX recombinant production of APAO polypeptides; as selection markers for
XX plant transformation; and to isolate related sequences from other
XX organisms. The APAO polypeptides are used to degrade mycotoxins in
XX plant materials, including expression in engineered bacteria and fungi,
XX e.g. rumen microflora.
XX
XX Sequence 3618 BP; 884 A; 925 C; 957 G; 852 T; 0 other:

Query Match 62.2%; Score 682.2; DB 21; Length 3618;
Best Local Similarity 99.6%; Pred. No. 56-178;
Matches 684; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ATGTCCTTACTAGTATTGGAATAAAGGCGCTTGCAACCCACTGACTTCTT 60
DB 1 atgtccctactactagttatttggaataataaggccttgcaaccactgactctt 60
QY 61 TTGGAATATCTTGAGAAAATATAGAGCATTTGTATGACCGGATGAAGTATAA 120
DB 61 ttggaatattcttgagaaaatataagagcatttgtagagcgagtagaagtgataaa 120
QY 121 TGGCAAAACAAAGATTGGAATTGGATTGGATTTCCCAATCTTCTTATTATTGAT 180
DB 121 tggcaaaacaaaagatttggaattggattggatttcccaatcttcttattatttgat 180
QY 181 GGTGATGTTAAATTAACACAGTCTATGCCATCATACGTTATATAGCTGACACAAAC 240
DB 181 ggtgatgtttaaattaacacagtctatgscatcatagctatataagctgacaagacaac 240

```

DB 181 ggtgacgtgtaaacatlaaacagctatgagccatcatatagctatataagctgacaaagcacaac 240
QY 241 ATGTGGGTGGTGTGTGTCACAAAGAGCGTGACAGATTTCATGCTGACAGAGGGGTTTG 300
DB 241 atgtgggtgggtgtgtgtcacaagaagcgtgcagagattccaatgctgaaagagcggtttg 300
QY 301 GATATTAGATACGGTGTTCGAGAAATGATATAGTAAAGACTTTGAACTGTCAAAATT 360
DB 301 gatattagatagcgtgttccgagaaattgacatataagaaagacttgaactctcaaaagt 360
QY 361 GATTTTCTAGACAGTACTGAAATGCTGAAATGTTGGAAGATCGTTATATGCTATAA 420
DB 361 gatcttcttagaagctaccctgaagaaatgctgaaatgctgaaatgcttataataa 420
QY 421 ACATATTAAATGCTGATCATGATTAACCCATCCGACTTCATGTTGATGAGCGCTTGAT 480
DB 421 acatattaaatgctgatacatgataaaccatcctgacttaattgtatagcgtcttgat 480
QY 481 GTTGTTTTATACATGAGCCCAATGTGCTGATGCTGCCAATTAAGTTGTTTAA 540
DB 481 gtgttttatacatgagcccaatgtgctgagtgctgccaaatagttgttttaa 540
QY 541 AAACGATTTGAGCTATCCCAAAATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 541 aaacgatattgagctatcccaaaaatgataagctgataatgataagctatataa 600
QY 601 TGGCCTTTGACAGGCTGGCAGACGATGTTGGTGGGAGACCATCCCAAAATCGAT 660
DB 601 tggcctttgacagcgtgagcaagcagcttgggtgggagaccatccccaatcgat 660
QY 661 CTGGTTCCGCGTGATCCCGAATTC 687
DB 661 ctggtccgcgtgagatcccggaattc 687

RESULT 13
ID 086774
ID 086774 standard; DNA: 699 BP.
AC 086774;
XX
DT 16-OCT-1995 (first entry)
XX
DE Glutathione-s-transferase gene.
XX
KW SDI-1; senescent cell-derived inhibitor; DNA synthesis;
KW senescence; cell proliferation; cancer; therapeutic; vulnerability;
KW fusion protein; glutathione-s-transferase; ss.
XX
OS Schistosoma japonicum.
XX
FH Key Location/Qualifiers
FT CDS 1..699
FT /tag= a
XX
PN W09506415-A.
XX
PD 09-MAR-1995.
XX
PE 26-AUG-1994; 94WO-US09700.
XX
PR 13-JUL-1994; 94US-0274535.
PR 30-AUG-1993; 93US-0113372.
PR 17-NOV-1993; 93US-0153564.
PR 03-JAN-1994; 94US-0160814.
PR 25-FEB-1994; 94US-0203535.
PR 15-APR-1994; 94US-0229420.
PR 30-JUN-1994; 94US-0268439.
XX
PA (BAYU) BAYLOR COLLEGE MEDICINE.
XX
XX Smith JR.
XX

DR WPI: 1995-131101/17.
DR P-PSDB: R72793.
XX
PT Nucleic acid encoding a protein or polypeptide that inhibits DNA
PT synthesis in a recipient cell - useful to inhibit cell
PT proliferation in tumour cells, treat wound or burn tissue, or as
PT an antiviral or antiparasitic agent
XX
PS Disclosure: Page 129; 169pp; English.
XX
CC The senescent cell-derived inhibitor-1 (SDI-1) inhibits DNA
CC synthesis in a recipient cell and is capable of associating with a
CC cyclin (esp. cyclin D1) or cyclin-dependent kinase (esp. CDK2).
CC Fusion of SDI-1 cDNA (086773) with a gene encoding S. japonicum
CC glutathione-s-transferase (086774) allowed production of a fusion
CC protein from which recombinant SDI-1 was easily isolated.
XX
SQ Sequence 699 BP; 205 A; 122 C; 156 G; 216 T; 0 other;

Query Match 62.18; Score 681; DB 15; Length 699;
Best Local Similarity 100.0%; Pred. No. 5.2e-178;
Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCCCCCTATACAGTATTGGAATTAAGGCGCTTGCAACCCACGACTTCTT 60
DB 1 atgccccctatacagctatttggaataaagagccttgcaaccacgcgacttctt 60
QY 61 TTGCAATATCTTGAAGAAAATATGAAGAGCATTTGTATGAGCCGATGAGTGTATAA 120
DB 61 ttgcaatattcttgaagaaaatatagaagcatttgatagcgcgatgaagtgataaa 120
QY 121 TGGGGAACAAAGATTGATTTGGCTTTGGCATTTCCCATCTTCTTATTAATTGAT 180
DB 121 tggggaacaaagatttgatttggctttggcatttcccatcttcttattatgat 180
QY 181 GGTGATGTTAAATTAACACAGTCTATGCGCATCATGATTATATAGTGACACACAAAC 240
DB 181 ggtgattgtaatttaacacagctctatgcgcatcatgattatattagtgacacacaac 240
QY 241 ATGTGGGTGGTGTTCACAAAGAGCGTGACAGATTTCAATGCTTGAAGAGCGGTTTG 300
DB 241 atgtgggtgggtgtgtcacaagaagcgtgcagagattccaatgcttgaagagcggtttg 300
QY 301 GATATTAGATACGGTGTTCGAGAAATGATATAGTAAAGACTTTGAACTGTCAAAATT 360
DB 301 gatattagatagcgtgttccgagaaattgacatataagaaagacttgaactctcaaaagt 360
QY 361 GATTTTCTAGACAGTACTGAAATGCTGAAATGTTGGAAGATCGTTATATGCTATAA 420
DB 361 gatcttcttagaagctaccctgaagaaatgctgaaatgctgaaatgcttataataa 420
QY 421 ACATATTAAATGCTGATCATGATTAACCCATCCGACTTCATGTTGATGAGCGCTTGAT 480
DB 421 acatattaaatgctgatacatgataaaccatcctgacttaattgtatagcgtcttgat 480
QY 481 GTTGTTTTATACATGAGCCCAATGTGCTGATGCTGCCAATTAAGTTGTTTAA 540
DB 481 gtgttttatacatgagcccaatgtgctgagtgctgccaaatagttgttttaa 540
QY 541 AAACGATTTGAGCTATCCCAAAATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 541 aaacgatattgagctatcccaaaaatgataagctgataatgataagctatataa 600
QY 601 TGGCCTTTGACAGGCTGGCAGACGATGTTGGTGGGAGACCATCCCAAAATCGAT 660
DB 601 tggcctttgacagcgtgagcaagcagcttgggtgggagaccatccccaatcgat 660
QY 661 CTGGTTCCGCGTGATCCCG 681
DB 661 ctggtccgcgtgagatcccg 681

Sequence	BP	A	328 C	497 G	308 T	0 other
Sequence 1695 BP:	562 A;	328 C;	497 G;	308 T;	0 other;	

Query Match	62.1%	Score 681	DB 16	Length 1695
Best Local Similarity	100.0%	Pred. NO. 7.7e-178		
Matches 681	Conservative	0	Mismatches 0	Indels 0
				Gaps 0
Qy	1	ANGCCCTATCTAGTATTGGAANAATTAAGGGCGCTTGTCCAAACCCACTGACTTCTT	60	
Db	1	agtcgccatctatctaggtatcttgaaanaatlaagggccttgcbaacccctcgaacttct	60	

OY	61	TTGAAATATCTTGAAGAAAAATATGAAGACGATTTGATATAGGCGCATGGAAGTGATAA	120
Db	61	ttggaatactcttgaagaanaaaatacgaagagattcttgatacgaacggaatgaagattgataa	120
OY	121	TGGCGAACAACAAAAGTTTGATTTGGATTTGGATTTGGAGTTTGGATTTCCCAATCTCTTATTTATATGAT	180
Db	121	tggcgaacaacaaaagtttgaatttggaatttggaatttggaatttcccaatcttcttatatattgat	180
OY	181	GGTGATGTTTAAATTAACACACTATGAGCCATCATTACGTTATATATAGCTGACAAGCACAC	240
Db	181	ggtgatgtttaaattaacacactatgagccatcattacgttatatatagctgacaaagcaaac	240
OY	241	ATGTTGGGTGGTGTGCCAAAAGACCGTGCGAGAGTTTCATGCTTGAAGAGCGGTTTTG	300
Db	241	atgttggttggtgtgtccaaaagaccgtgcgagagatttcaatgcttgaagagcggttttg	300
OY	301	GATATTACATACGGTGTTTCCAGAAATTGACATATATGATAAAGACTTTGAAACCTCCAAAGTT	360
Db	301	gatatltgaatacgtgtgttctgagagattgcataatagtaagaacttgaacttccaaagtc	360
OY	361	GATTTTCTTACCAAGCTACCTGAAGTCGTAAGAAVTGTTCCAAGATCGTTTATGTCAATAA	420
Db	361	gatttcttctaacgaacgtccggaagtgcgtgaagaatgttcgaagatctgtatgtlcataaa	420
OY	421	ACATATTTAAATGGGATCATGTAAACCATTCCTACCTTCATGTTGTATGAGCGCTTGAT	480
Db	421	acatatcttaaatggtgatcatgtatacccaatccctgaacttcaatgtgtatagtcgccttgat	480
OY	481	GTTGTTTATACATGAGCCCAATGTGCGTGAGTCGTTCCCAAAATTAGTTTGTTTTAA	540
Db	481	gttgtttatcatatlgagcccaatgtgcctgagatgcttcccaaatatgatttgtttaa	540
OY	541	AAACGTATTGAAGCATGCCCAAAATTTGATATGATACCTTGAATCCACACAGTATATGCA	600
Db	541	aaacgtatttgaagcatgtcccaaattgataagtaacttgaatccacagcaaglatatagca	600
OY	601	TGGCCTTTGACAGGGCTGCAAGCACGTTTGTTGTTGGCGACATCTCCAAAATCGAT	660
Db	601	tggcctttgacagggctgcgaagcacgtttggttggttggcgacatctctccaaaatcgat	660
OY	661	CTMGTTCCGCGTGATCCCG 661	
Db	661	ctmgttccgcgtgatcccg 661	

RESULT	15
ID	T05332
XX	T05332 standard; DNA; 1932 BP.
XX	
AC	T05332;
XX	
DT	09-MAY-1996 (first entry)
XX	
DE	DNA encoding TCR27 polypeptide, Ag44.
XX	
'KW	TCR27; Chagas disease; repeat unit; diagnosis; blood screening;
XX	recombinant; fusion protein; glutathione-S-transferase; ds.
XX	
OS	Trypanosoma cruzi.
XX	
XX	
FH	Key
FT	CDS
FT	
FT	Location/Qualifiers
FT	1..1932
FT	/*tag a
FT	/product= Ag44
FT	/note= "Gst-TCR27-Ag15-fusion protein"
FT	985..1656
FT	/*tag= b
FT	/note= "encodes 16 of 69 repeat units of 14 amino
XX	acids"
XX	
PN	W09525797-A1.
XX	

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 18, 2001, 03:47:41 ; Search time 56.88 Seconds

(without alignments)
3105.339 Million cell updates/sec

Title: US-09-402-488A-1

Perfect score: 1096
Sequence: 1 ATGTCCCTACTAGTGA.....ATCCACTAGTTAGACGG 1096

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 280836 seqs, 80580151 residues

Total number of hits satisfying chosen parameters: 561672

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA.*

1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/PCRB_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	681	62.1	699	4	PCT-US94-09700-5
2	681	62.1	1065	2	US-08-216-894-3
3	681	62.1	1695	2	US-08-216-894-1
4	681	62.1	1932	2	US-08-216-894-7
5	680	62.0	1194	4	PCT-US94-09700-10
6	678.4	61.9	1281	3	US-08-737-248-3
7	678.2	61.9	1740	2	US-08-864-224-10
8	667.4	60.9	897	1	US-08-102-757-8
9	663.4	60.5	1419	2	US-08-216-894-9
10	663.4	60.5	1920	3	US-08-974-180-19
11	662.8	60.5	2095	2	US-08-305-764C-55
12	661	60.3	924	2	US-08-216-894-5
13	660.4	60.3	1951	3	US-08-895-707-1
14	654	59.7	654	1	US-08-102-757-25
15	649	59.2	897	1	US-08-102-757-10
16	151.8	13.9	1458	1	US-07-854-596B-42
17	151.8	13.9	1467	1	US-07-854-596B-46
18	151.6	13.8	201	1	US-07-854-596B-1
19	151.6	13.8	223	1	US-07-854-596B-7
20	151.6	13.8	420	1	US-07-854-596B-8
21	144.2	13.2	1082	1	US-08-715-252-1
22	144.2	13.2	1082	2	US-08-715-252-1
23	144.2	13.2	1130	1	US-08-453-051-3
24	140.6	12.8	227	5	5164304-3
25	140.2	12.8	217	5	5422249-13
26	139.4	12.7	563	1	US-08-044-506B-22
27	138.6	12.6	206	1	US-08-044-506B-21
28	138	12.6	195	1	US-08-406-948A-7

29	137.8	12.6	279	1	US-08-186-222-3	Sequence 3, Appl
30	137.6	12.6	195	1	US-08-044-506B-27	Sequence 27, Appl
31	137	12.5	202	1	US-08-044-506B-20	Sequence 20, Appl
32	136.4	12.4	195	1	US-08-044-506B-28	Sequence 28, Appl
33	131.2	12.0	212	5	5180668-10	Patent No. 5180668
34	131.2	12.0	227	2	US-07-982-064-8	Sequence 8, Appl
35	131.2	12.0	8491	2	US-08-757-439-1	Sequence 1, Appl
36	129.8	11.8	2115	2	US-08-767-026-3	Sequence 3, Appl
37	129	11.8	1240	1	US-08-240-372-2	Sequence 2, Appl
38	127.6	11.6	218	5	5422249-4	Patent No. 5422249
39	127	11.6	2732	5	5217891-14	Patent No. 5217891
40	127	11.6	2733	2	US-08-846-021A-6	Sequence 6, Appl
41	126.8	11.6	235	1	US-08-262-384A-7	Sequence 7, Appl
42	126.8	11.6	235	1	US-08-262-384A-8	Sequence 8, Appl
43	126.8	11.6	726	1	US-08-262-384A-9	Sequence 9, Appl
44	126.8	11.6	726	1	US-08-262-384A-10	Sequence 10, Appl
45	125.8	11.5	198	1	US-07-910-528-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
PCT-US94-09700-5
Sequence 5, Application PC/TUS9409700
GENERAL INFORMATION:
APPLICANT: BAYLOR COLLEGE OF MEDICINE
APPLICANT: SMITH, JAMES R.
TITLE OF INVENTION: SENESCENT CELL DERIVED INHIBITORS OF
TITLE OF INVENTION: DNA SYNTHESIS
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESSES:
ADDRESSEE: HOMREY & SIMON
STREET: 1299 PENNSYLVANIA AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/09700
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/808,523
FILING DATE: 16-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/970,462
FILING DATE: 02-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/113,372
FILING DATE: 30-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/153,564
FILING DATE: 17-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/203,535
FILING DATE: 25-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/229,420
FILING DATE: 15-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/274,535
FILING DATE: 13-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: AUERBACH, JEFFREY I.
REGISTRATION NUMBER: 32,680
REFERENCE/DOCKET NUMBER: 225-102-CIP7-PCT
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 383-7451
TELEFAX: (202) 383-6610
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 699 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Schistosoma japonicum
IMMEDIATE SOURCE:
CLONE: GST
PCT-US94-09700-5

Query Match 62.1%; Score 681; DB 4; Length 699;
Best Local Similarity 100.0%; Pred. No. 9,3e-185;
Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 ATGTCCTATACACTGATGTTATGGAATTAAGGCTTGTCACCACTGACTTCTT 60
    |||||||
DB 1 ATGTCCTATACACTGATGTTATGGAATTAAGGCTTGTCACCACTGACTTCTT 60

OY 61 TTGGAATCTTGGAAGAAAATGAGAGCATTTGTATGAGCGGATGAAGTATAA 120
    |||||||
DB 61 TTGGAATCTTGGAAGAAAATGAGAGCATTTGTATGAGCGGATGAAGTATAA 120

OY 121 TGGCAAAACAAATTTGAATTTGGTTTCCCAATCTTCTATATATATGAT 180
    |||||||
DB 121 TGGCAAAACAAATTTGAATTTGGTTTCCCAATCTTCTATATATGAT 180

OY 181 TGGCAAAACAAATTTGAATTTGGTTTCCCAATCTTCTATATATATGAT 240
    |||||||
DB 181 TGGCAAAACAAATTTGAATTTGGTTTCCCAATCTTCTATATATGAT 240

OY 241 ATGTTGGGGGTGGTTCACAAAGAGGTGAGATTTCAATCTTGAAGGAGGTTTG 300
    |||||||
DB 241 ATGTTGGGGGTGGTTCACAAAGAGGTGAGATTTCAATCTTGAAGGAGGTTTG 300

OY 301 GATATTAGATACGTTGTTGAGAAATTCATATAGTAAGACTTTGAAATCTCAAGTT 360
    |||||||
DB 301 GATATTAGATACGTTGTTGAGAAATTCATATAGTAAGACTTTGAAATCTCAAGTT 360

OY 361 GATTTTCTTAGACACCTACCTGAAATGCTGAAATGTTGGAAGATCGTTATGATATA 420
    |||||||
DB 361 GATTTTCTTAGACACCTACCTGAAATGCTGAAATGTTGGAAGATCGTTATGATATA 420

OY 421 ACATATTTAAATGATGATCATGTAACCCATCTGACTCATGTTGATGACCTTTGAT 480
    |||||||
DB 421 ACATATTTAAATGATGATCATGTAACCCATCTGACTCATGTTGATGACCTTTGAT 480

OY 481 GTTGTGTTTATACATGACCAATGTCCTGATGCGTTCCCAAAATTAATTTGTTTAA 540
    |||||||
DB 481 GTTGTGTTTATACATGACCAATGTCCTGATGCGTTCCCAAAATTAATTTGTTTAA 540

OY 541 AAACGATTTGAAGCTATCCCAAAATGTAAGTACTTAATTCACGCAAGATATAGCA 600
    |||||||
DB 541 AAACGATTTGAAGCTATCCCAAAATGTAAGTACTTAATTCACGCAAGATATAGCA 600

OY 601 TGGCCCTTTCAGAGGTGGGAGAGCAAGTTGTTGTTGGGAGACATCTCCAAATTCGAT 660
    |||||||
DB 601 TGGCCCTTTCAGAGGTGGGAGAGCAAGTTGTTGTTGGGAGACATCTCCAAATTCGAT 660

OY 661 CTGGTTCGCGGTGATCCCG 681
    |||||||
DB 661 CTGGTTCGCGGTGATCCCG 681
```

RESULT 2
US-08-216-894-3

Sequence 3, Application US/08216894
Patent No. 5876734
GENERAL INFORMATION:
APPLICANT: Kitchhoff, Louis V.
APPLICANT: Otsu, Keiko
TITLE OF INVENTION: POLYPEPTIDES FOR DIAGNOSING INFECTION
TITLE OF INVENTION: WITH TRYPAOSOMA CRUZI
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/216,894
FILING DATE: 24-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 85326/102/DRL0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1065 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1062
US-08-216-894-3

Query Match 62.1%; Score 681; DB 2; Length 1065;
Best Local Similarity 100.0%; Pred. No. 1.1e-184;
Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 ATGTCCTATACACTGATGTTATGGAATTAAGGCTTGTCACCACTGACTTCTT 60
    |||||||
DB 1 ATGTCCTATACACTGATGTTATGGAATTAAGGCTTGTCACCACTGACTTCTT 60

OY 61 TTGGAATCTTGGAAGAAAATGAGAGCATTTGTATGAGCGGATGAAGTATAA 120
    |||||||
DB 61 TTGGAATCTTGGAAGAAAATGAGAGCATTTGTATGAGCGGATGAAGTATAA 120

OY 121 TGGCAAAACAAATTTGAATTTGGTTTCCCAATCTTCTATATATATGAT 180
    |||||||
DB 121 TGGCAAAACAAATTTGAATTTGGTTTCCCAATCTTCTATATATATGAT 180

OY 181 GGTGATGTTAAATTAACACAGCTATGAGCAATCATAGCTTATATGCTGACAGCACAC 240
    |||||||
DB 181 GGTGATGTTAAATTAACACAGCTATGAGCAATCATAGCTTATATGCTGACAGCACAC 240

OY 241 ATGTTGGGGTGGTTCACAAAGAGGTGAGATTTCAATGCTTTGAAGAGGCGTTTG 300
    |||||||
DB 241 ATGTTGGGGTGGTTCACAAAGAGGTGAGATTTCAATGCTTTGAAGAGGCGTTTG 300

OY 301 GATATTAGATACGTTGTTGCAAGATTCGATATAGTAAGACTTTGAAATCTCAAGTT 360
    |||||||
DB 301 GATATTAGATACGTTGTTGCAAGATTCGATATAGTAAGACTTTGAAATCTCAAGTT 360

OY 361 GATTTTCTTAGACACCTGAAATGCTGAAATGTTGGAAGATCGTTATGATATA 420
    |||||||
```


Db 361 GATTTCTTAGACAGCTACCTGAATGCTGAAATGTGCAAGATCGTTATGTCTATAA 420
Qy 421 ACATATTTAAATGATGATCATGTAACCCATCCGACTTCATGTTATGACGCTTGAT 480
Db 421 ACATATTTAAATGATGATCATGTAACCCATCCGACTTCATGTTATGACGCTTGAT 480
Qy 481 GTTGTTTTACATGAGCAACCAATGTGCTGATGCTGCTCCAAATTAATGTTTAA 540
Db 481 GTTGTTTTACATGAGCAACCAATGTGCTGATGCTGCTCCAAATTAATGTTTAA 540
Qy 541 AAACGATTTGAAGCTATGCCACAATTAATGATTAATGATTAATGATTAATGATTA 600
Db 541 AAACGATTTGAAGCTATGCCACAATTAATGATTAATGATTAATGATTAATGATTA 600
Qy 601 TGCGCTTTGACGAGGCTGGCAAGCAGCTTTGTTGTTGAGGACCATCTCCAAATGAGAT 660
Db 601 TGCGCTTTGACGAGGCTGGCAAGCAGCTTTGTTGTTGAGGACCATCTCCAAATGAGAT 660
Qy 661 CTGGTTCCCGCTGATCCCG 681
Db 661 CTGGTTCCCGCTGATCCCG 681

RESULT 3

US-08-216-894-1
; Sequence 1, Application US/08216894
; Patent No. 5876734
; GENERAL INFORMATION:
; APPLICANT: Kirchoff, Louis V.
; APPLICANT: Otsu, Keiko
; TITLE OF INVENTION: POLYPEPTIDES FOR DIAGNOSING INFECTION
; TITLE OF INVENTION: WITH TRYPAOSOMA CRUZI
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/216,894
; FILING DATE: 24-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bent, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 85326/102/DRLO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1695 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1692
; US-08-216-894-1

Query Match 62.1%; Score 681; DB 2; Length 1695;
Best Local Similarity 100.0%; Pred. No. 1.4e-184;
Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTCCCTATACCTAGCTATTGGAATAATTAAGGCGCTTGTCAGACCACTCGACTTCTT 60

Db 1 ATGTCCCTATACCTAGCTATTGGAATAATTAAGGCGCTTGTCAGACCACTCGACTTCTT 60
Qy 61 TTGGATATCTGGAAGAAAATATGAAGACATTTGTATGAGCGGATGAAGTGATTA 120
Db 61 TTGGATATCTGGAAGAAAATATGAAGACATTTGTATGAGCGGATGAAGTGATTA 120
Qy 121 TGCGCAACAAAAGTTGAATTTGGTTTGAAGTTTCCCAATCTTATTAATTAAT 180
Db 121 TGCGCAACAAAAGTTGAATTTGGTTTGAAGTTTCCCAATCTTATTAATTAAT 180
Qy 181 GTGATGTTAATTAACACAGTCTATGCGCATATACGTTATATAGCTGACAGCAAC 240
Db 181 GTGATGTTAATTAACACAGTCTATGCGCATATACGTTATATAGCTGACAGCAAC 240
Qy 241 ATGTTGGTGTTGTTGTCAAAAGAGCGTGAGATTTCAATGCTTGAGAGCGGTTTG 300
Db 241 ATGTTGGTGTTGTTGTCAAAAGAGCGTGAGATTTCAATGCTTGAGAGCGGTTTG 300
Qy 301 GATATTAGATACGCTGTTTCGAGAAATGATATAGTAAGACTTTGAACCTCAAGTT 360
Db 301 GATATTAGATACGCTGTTTCGAGAAATGATATAGTAAGACTTTGAACCTCAAGTT 360
Qy 361 GATTTCTTAGACAGCTACCTGAATGCTGAATGTTGAAGATGTTATGTCTATAA 420
Db 361 GATTTCTTAGACAGCTACCTGAATGCTGAATGTTGAAGATGTTATGTCTATAA 420
Qy 421 ACATATTTAAATGATGATATGTAACCCATCCGACTTCATGTTATGATGCTTGAT 480
Db 421 ACATATTTAAATGATGATATGTAACCCATCCGACTTCATGTTATGATGCTTGAT 480
Qy 481 GTTGTTTTACATGAGCAACCAATGCTGATGCTGCTCCCAAAATTAATGTTTAA 540
Db 481 GTTGTTTTACATGAGCAACCAATGCTGATGCTGCTCCCAAAATTAATGTTTAA 540
Qy 541 AAACGATTTGAAGCTATGCCACAATTAATGATTAATGATTAATGATTAATGATTA 600
Db 541 AAACGATTTGAAGCTATGCCACAATTAATGATTAATGATTAATGATTAATGATTA 600
Qy 601 TGCGCTTTGACGAGGCTGGCAAGCAGCTTTGTTGTTGAGGACCATCTCCAAATGAGAT 660
Db 601 TGCGCTTTGACGAGGCTGGCAAGCAGCTTTGTTGTTGAGGACCATCTCCAAATGAGAT 660
Qy 661 CTGGTTCCCGCTGATCCCG 681
Db 661 CTGGTTCCCGCTGATCCCG 681

RESULT 4

US-08-216-894-7
; Sequence 7, Application US/08216894
; Patent No. 5876734
; GENERAL INFORMATION:
; APPLICANT: Kirchoff, Louis V.
; APPLICANT: Otsu, Keiko
; TITLE OF INVENTION: POLYPEPTIDES FOR DIAGNOSING INFECTION
; TITLE OF INVENTION: WITH TRYPAOSOMA CRUZI
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/216,894
; FILING DATE: 24-MAR-1994
; ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 85326/102/DRLO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1932 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1929
US-08-216-894-7

Query Match 62.1%; Score 681; DB 2; Length 1932;

Best Local Similarity 100.0%; Pred. No. 1.5e-184; Mismatches 0; Indels 0; Gaps 0;

Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTCCTATACGTTAGTAAATTAAGGCTTGTCACACCACTGACTTCT 60
DB 1 ATGTCCTATACGTTAGTAAATTAAGGCTTGTCACACCACTGACTTCT 60
QY 61 TTGGAATCTTGAAGAAATATAGAGCATTTGATGAGCGGATGAGGTATAA 120
DB 61 TTGGAATCTTGAAGAAATATAGAGCATTTGATGAGCGGATGAGGTATAA 120
QY 121 TGGCGAACAATAATTTGAATGGGTTGGATTTCCCAATCTCTATATATGAT 180
DB 121 TGGCGAACAATAATTTGAATGGGTTGGATTTCCCAATCTCTATATATGAT 180
QY 181 GGTGATGTTAAATTAACACAGTCTATGSCCATACATAGCTTATATAGCTGACAAGCACAA 240
DB 181 GGTGATGTTAAATTAACACAGTCTATGSCCATACATAGCTTATATAGCTGACAAGCACAA 240
QY 241 ATGTTGGGTTGGTTCACAAAGAGGCTGACAGATTTCAATGCTTGAAGAGCGGTTTG 300
DB 241 ATGTTGGGTTGGTTCACAAAGAGGCTGACAGATTTCAATGCTTGAAGAGCGGTTTG 300
QY 301 CATATTAGATACGCTGTTGAGAAATGCAATAGTAAAGACTTTGAACTCTCAAAGT 360
DB 301 CATATTAGATACGCTGTTGAGAAATGCAATAGTAAAGACTTTGAACTCTCAAAGT 360
QY 361 GATTTTCTTAGCAGCTACCTGAAATGCTGAAATGTTGGAAGATGTTATGTCATAAA 420
DB 361 GATTTTCTTAGCAGCTACCTGAAATGCTGAAATGTTGGAAGATGTTATGTCATAAA 420
QY 421 ACATTTAAATGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
DB 421 ACATTTAAATGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 481 GTTGTTTTATACATGACCAATGCTGATGATGATGATGATGATGATGATGATGATGAT 540
DB 481 GTTGTTTTATACATGACCAATGCTGATGATGATGATGATGATGATGATGATGATGAT 540
QY 541 AAACGATTTGAAGCTATCCACAAATTTGATAGTACTTGAATTCACAGCAATATAGCA 600
DB 541 AAACGATTTGAAGCTATCCACAAATTTGATAGTACTTGAATTCACAGCAATATAGCA 600
QY 601 TGGCCTTTGACAGGCTGACAGCAAGTTGTTGGTGGGAGACATCTCCAAATGCGAT 660
DB 601 TGGCCTTTGACAGGCTGACAGCAAGTTGTTGGTGGGAGACATCTCCAAATGCGAT 660
QY 661 CTGTTCCGCGTGAATCCCG 681
DB 661 CTGTTCCGCGTGAATCCCG 681

RESULT 5
PCT-US94-09700-10
Sequence 10. Application PC/TUS9409700
GENERAL INFORMATION:
APPLICANT: BAYLOR COLLEGE OF MEDICINE
APPLICANT: SMITH, JAMES R.
TITLE OF INVENTION: SENSITIVE CELL DERIVED INHIBITORS OF
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOWREY & SIMON
STREET: 1299 PENNSYLVANIA AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/09700
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/808,523
FILING DATE: 16-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/970,462
FILING DATE: 02-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/113,372
FILING DATE: 30-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/153,564
FILING DATE: 17-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/203,535
FILING DATE: 25-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/229,420
FILING DATE: 15-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/274,535
FILING DATE: 13-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: AUBRECHT, JEFFREY I.
REGISTRATION NUMBER: 32,680
REFERENCE/DOCKET NUMBER: 225-102-CIP7-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 383-7451
TELEFAX: (202) 383-6610
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1194 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTISENSE: NO
IMMEDIATE SOURCE:
CLONE: GST-SDI-1 gene fusion
PCT-US94-09700-10

Query Match 62.0%; Score 680; DB 4; Length 1194;
Best Local Similarity 100.0%; Pred. No. 2.3e-184; Mismatches 0; Indels 0; Gaps 0;
Matches 680; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTCCTATACGTTAGTAAATTAAGGCTTGTCACACCACTGACTTCT 60
DB 1 ATGTCCTATACGTTAGTAAATTAAGGCTTGTCACACCACTGACTTCT 60

```

Db 1 ATGTCCCTATACGTATTTGAAATTTAGGGCTGTGCAACCCACTGACTTCT 60
Qy 61 TTGGATATCTTGAAGAAAAATATGAGAGCACTTTGTATGAGCGGATGAGTATAA 120
Db 61 TTGGATATCTTGAAGAAAAATATGAGAGCACTTTGTATGAGCGGATGAGTATAA 120
Qy 121 TGGCGAACAACAAAGTTGAATTTGGSTTTGAGTTTCCCAATCTTCTTATTTATGAT 180
Db 121 TGGCGAACAACAAAGTTGAATTTGGSTTTGAGTTTCCCAATCTTCTTATTTATGAT 180
Qy 181 GGTGATGTTAAATTAACACAGTCTATGAGCCATCATCTTATATATGCTGACAGCAAC 240
Db 181 GGTGATGTTAAATTAACACAGTCTATGAGCCATCATCTTATATATGCTGACAGCAAC 240
Qy 241 ATGTGGGTGGTTGTCCAAAGAGCGTGAGATTTCAATGCTTGAAGAGCGGTTTGG 300
Db 241 ATGTGGGTGGTTGTCCAAAGAGCGTGAGATTTCAATGCTTGAAGAGCGGTTTGG 300
Qy 301 GATATTAGTACGGGTTCGAGAAATGATATAGTAAAGACTTTGAAACTCTCAAGTT 360
Db 301 GATATTAGTACGGGTTCGAGAAATGATATAGTAAAGACTTTGAAACTCTCAAGTT 360
Qy 361 GATTTCTTAGCAGCTACCTGAATGCTGAAATGTTGAGAGATGCTTATGCTTGAAT 420
Db 361 GATTTCTTAGCAGCTACCTGAATGCTGAAATGTTGAGAGATGCTTATGCTTGAAT 420
Qy 421 ACATATTTAAATGATGATCATGTATACCATCTGACTTCATGTTATGAGCGCTTTGAT 480
Db 421 ACATATTTAAATGATGATCATGTATACCATCTGACTTCATGTTATGAGCGCTTTGAT 480
Qy 481 GTTGTTTTATACATGAGCAATGTGCTGATGCTTCCCAAAATTTAGTTTGTAAAA 540
Db 481 GTTGTTTTATACATGAGCAATGTGCTGATGCTTCCCAAAATTTAGTTTGTAAAA 540
Qy 541 AAACGATTGAAGCTATCCACAAATTTGATAGTACTTGAATCCAGCAATATAGCA 600
Db 541 AAACGATTGAAGCTATCCACAAATTTGATAGTACTTGAATCCAGCAATATAGCA 600
Qy 601 TGGCCCTTTCAGAGGCTGAGCAAGCCAGCTTGTGTGTGGGACCATCTCCAAATCGAT 660
Db 601 TGGCCCTTTCAGAGGCTGAGCAAGCCAGCTTGTGTGTGGGACCATCTCCAAATCGAT 660
Qy 661 CTGGTCCCGGTGATCCCC 680
Db 661 CTGGTCCCGGTGATCCCC 680

```

RESULT 6
US-08-737-248-3
Sequence 3, Application US/08737248
Patent No. 6114305

GENERAL INFORMATION:
APPLICANT: Guemene, Daniel
APPLICANT: zadworny, David
APPLICANT: Karalzas, Costas
TITLE OF INVENTION: USE OF PROLACTIN FOR PREVENTING AND/OR
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: WEISER & ASSOCIATES
STREET: 230 South Fifteenth Street, Suite 500
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19102
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/737,248
FILING DATE: 28-APR-1997

```

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR95/00576
FILING DATE: 03-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 94/05550
FILING DATE: 05-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Weiser, Gerard J.
REGISTRATION NUMBER: 19,763
REFERENCE/DOCKET NUMBER: 989,6411P
TELEPHONE: 215-875-8394
TELEFAX: 215-875-8394
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1281 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1278
US-08-737-248-3

```

Query Match 61.9%; Score 678.4; DB 3; Length 1281;
Best Local Similarity 99.9%; Pred. No. 6.8e-184;
Matches 679; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy 1 ATGTCCCTATACGTATTTGAAATTTAGGGCTGTGCAACCCACTGACTTCT 60
Db 1 ATGTCCCTATACGTATTTGAAATTTAGGGCTGTGCAACCCACTGACTTCT 60
Qy 61 TTGGATATCTTGAAGAAAAATATGAGAGCACTTTGTATGAGCGGATGAGTATAA 120
Db 61 TTGGATATCTTGAAGAAAAATATGAGAGCACTTTGTATGAGCGGATGAGTATAA 120
Qy 121 TGGCGAACAACAAAGTTGAATTTGGSTTTGAGTTTCCCAATCTTCTTATTTATGAT 180
Db 121 TGGCGAACAACAAAGTTGAATTTGGSTTTGAGTTTCCCAATCTTCTTATTTATGAT 180
Qy 181 GGTGATGTTAAATTAACACAGTCTATGAGCCATCATCTTATATATGCTGACAGCAAC 240
Db 181 GGTGATGTTAAATTAACACAGTCTATGAGCCATCATCTTATATATGCTGACAGCAAC 240
Qy 241 ATGTGGGTGGTTGTCCAAAGAGCGTGAGATTTCAATGCTTGAAGAGCGGTTTGG 300
Db 241 ATGTGGGTGGTTGTCCAAAGAGCGTGAGATTTCAATGCTTGAAGAGCGGTTTGG 300
Qy 301 GATATTAGTACGGGTTCGAGAAATGATATAGTAAAGACTTTGAAACTCTCAAGTT 360
Db 301 GATATTAGTACGGGTTCGAGAAATGATATAGTAAAGACTTTGAAACTCTCAAGTT 360
Qy 361 GATTTCTTAGCAGCTACCTGAATGCTGAAATGTTGAGAGATGCTTATGCTTGAAT 420
Db 361 GATTTCTTAGCAGCTACCTGAATGCTGAAATGTTGAGAGATGCTTATGCTTGAAT 420
Qy 421 ACATATTTAAATGATGATCATGTATACCATCTGACTTCATGTTATGAGCGCTTTGAT 480
Db 421 ACATATTTAAATGATGATCATGTATACCATCTGACTTCATGTTATGAGCGCTTTGAT 480
Qy 481 GTTGTTTTATACATGAGCAATGTGCTGATGCTTCCCAAAATTTAGTTTGTAAAA 540
Db 481 GTTGTTTTATACATGAGCAATGTGCTGATGCTTCCCAAAATTTAGTTTGTAAAA 540
Qy 541 AAACGATTGAAGCTATCCACAAATTTGATAGTACTTGAATCCAGCAATATAGCA 600
Db 541 AAACGATTGAAGCTATCCACAAATTTGATAGTACTTGAATCCAGCAATATAGCA 600
Qy 601 TGGCCCTTTCAGAGGCTGAGCAAGCCAGCTTGTGTGTGGGACCATCTCCAAATCGAT 660
Db 601 TGGCCCTTTCAGAGGCTGAGCAAGCCAGCTTGTGTGTGGGACCATCTCCAAATCGAT 660

```

Db 601 TGGCCTTTCAGGGCTGGCAAGCCAGCTTTGGTGGCGAGCAATCTCCAAATGGAT 660
QY 661 CTGGTTCCCGTGGATCCCC 680
|||||
Db 661 CTGGTTCCCGTGGATCCTC 680

RESULT 7

US-08-864-224-10

Sequence 10, Application US/08864224

Patent No. 5851808

GENERAL INFORMATION:

APPLICANT: Eileedge, Stephen J.

APPLICANT: Liu, Qinyhua

TITLE OF INVENTION: Rapid Subcloning Using Site-Specific

NUMBER OF SEQUENCES: 18

CORRESPONDENCE ADDRESS:

ADDRESSEE: Medlen & Carroll, LLP

STREET: 220 Montgomery Street, Suite 2200

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/864,224

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Ingolia, Diane E.

REGISTRATION NUMBER: 40,027

REFERENCE/DOCKET NUMBER: BCM-02681

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 705-8410

TELEFAX: (415) 397-8338

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 1740 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid

DESCRIPTION: /desc = "DNA"

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1737

US-08-864-224-10

Query Match

Best Local Similarity 98.1%; Score 678.2; DB 2; Length 1740;

Matches 686; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 ATGTCCCTATAGTATGTAATTAAGGCGCTTGTCACCCAGCTGCTTCT 60
|||||
Db 1 ATGTCCCTATAGTATGTAATTAAGGCGCTTGTCACCCAGCTGCTTCT 60
61 TTGGAATATCTTGAAGAAATATGAGCAATTTGTATGAGCGGATGAAGTATATA 120
|||||
Db 61 TTGGAATATCTTGAAGAAATATGAGCAATTTGTATGAGCGGATGAAGTATATA 120
121 TGGCAAAACAAAATTTGAATGGGTTTGAAGTTTCCCAATCTTCTATTATTGAT 180
|||||
Db 121 TGGCAAAACAAAATTTGAATGGGTTTGAAGTTTCCCAATCTTCTATTATTGAT 180
181 GGTATGTTAAATTAACAGCTATGGCCATCATAGCTTATATAGTGACAAAGCAAC 240
|||||
Db 181 GGTATGTTAAATTAACAGCTATGGCCATCATAGCTTATATAGTGACAAAGCAAC 240

QY 241 ATGTGGGTGGTGTTCCTCAAAAGAGCGTCAGATTTTCATGCTTGAAGAGCGGTTTG 300
|||||
Db 241 ATGTGGGTGGTGTTCCTCAAAAGAGCGTCAGATTTTCATGCTTGAAGAGCGGTTTG 300
301 GATATTAGATACGGTGTTCGAGAAATTCATATAGTAAAGACTTTGAACCTCTCAAGTT 360
|||||
Db 301 GATATTAGATACGGTGTTCGAGAAATTCATATAGTAAAGACTTTGAACCTCTCAAGTT 360
361 GATTTTCTAGCAACCTACCTGAATGCTGAATAATGCTGAAGATGTTTATGATATA 420
|||||
QY 361 GATTTTCTAGCAACCTACCTGAATGCTGAATAATGCTGAAGATGTTTATGATATA 420
421 GATTTTCTAGCAACCTACCTGAATGCTGAATAATGCTGAAGATGTTTATGATATA 420
Db 421 GATTTTCTAGCAACCTACCTGAATGCTGAATAATGCTGAAGATGTTTATGATATA 420
421 ACATATTTAAATGCTATCATATGTAACCAATCTGACTCATGTTTATGATAGTAT 480
Db 421 ACATATTTAAATGCTATCATATGTAACCAATCTGACTCATGTTTATGATAGTAT 480
481 GTTGTATATACATGAGACCAATGCTGCTGATGCTGCTCCCAAAATTTAGTTTATA 540
Db 481 GTTGTATATACATGAGACCAATGCTGCTGATGCTGCTCCCAAAATTTAGTTTATA 540
541 AAAGCTATTTGAAGCTATCCCAATTTGATAGTACTTGAATCCAGCAATATATAGCA 600
QY 541 AAAGCTATTTGAAGCTATCCCAATTTGATAGTACTTGAATCCAGCAATATATAGCA 600
601 TGGCCTTTCAGAGGCTGGCAAGCCAGCTTTGGTGGCGACCATCTCCAAATCGGAT 660
Db 601 TGGCCTTTCAGAGGCTGGCAAGCCAGCTTTGGTGGCGACCATCTCCAAATCGGAT 660
661 CTGGTTCCCGGTGATCCCCGAAATTCGCGGTCGACTG 699
QY 661 CTGGTTCCCGGTGATCCCCGAAATTCGCGGTCGACTG 699
Db 661 CTGGTTCCCGGTGATCCCCGAAATTCGCGGTCGACTG 699

RESULT 8

US-08-102-757-8

Sequence 8, Application US/08102757

Patent No. 5480864

GENERAL INFORMATION:

APPLICANT: Tajima, Masahiro

APPLICANT: Ohnuma, Mamami

APPLICANT: Lerner, Ethan A.

TITLE OF INVENTION: Modified Maxadilan Protein, Its

NUMBER OF SEQUENCES: 25

CORRESPONDENCE ADDRESS:

ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &

ADDRESS: Dunner

STREET: 1300 I Street, N.W., Suite 700

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/102,757

FILING DATE: 01-JAN-1993

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Meyers, Kenneth J.

REGISTRATION NUMBER: 25,146

REFERENCE/DOCKET NUMBER: 05136.0003-00000

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-408-4000

TELEFAX: 202-408-4400

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 897 base pairs

TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-102-757-8

Query Match 60.9%; Score 667.4; DB 1; Length 897;
Best Local Similarity 99.1%; Pred. No. 7.8e-181;
Matches 671; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 ATGCCCCCTATCTAGGTTATGGAATAATAGAGCCCTTGCCACCCACTGACTTCT 60
DB 1 ATGCCCCCTATCTAGGTTATGGAATAATAGAGCCCTTGCCACCCACTGACTTCT 60
OY 61 TTGGAATATCTTGAGAAATAATATGAGAGCATTTGATAGCCGAGTAAAGTATAA 120
DB 61 TTGGAATATCTTGAGAAATAATATGAGAGCATTTGATAGCCGAGTAAAGTATAA 120
OY 121 TGGCCAAACAAAGTTGATGAGTTGGGTTGAGCTTCCCAATCTCTTATATATGAT 180
DB 121 TGGCCAAACAAAGTTGATGAGTTGGGTTGAGCTTCCCAATCTCTTATATATGAT 180
OY 181 GGTGATGTTAAATTAACAGCTATAGCCATCATAGCTATATAGCTGACAGCAAC 240
DB 181 GGTGATGTTAAATTAACAGCTATAGCCATCATAGCTATATAGCTGACAGCAAC 240
OY 241 ATGTTGGGTGGTTGTCACAAAGAGCGTGAGAGATTTCAATGCTTGAAGAGCGTTTG 300
DB 241 ATGTTGGGTGGTTGTCACAAAGAGCGTGAGAGATTTCAATGCTTGAAGAGCGTTTG 300
OY 301 GATATTAGATACGTTGTTGCAAAATTCGATATAGTAAAGACTTTGAAACTCTCAAGTT 360
DB 301 GATATTAGATACGTTGTTGCAAAATTCGATATAGTAAAGACTTTGAAACTCTCAAGTT 360
OY 361 GATTTTCTTAGACACTACCTAATGCTGAAATGTTGGAAGATGTTATATGCTATAA 420
DB 361 GATTTTCTTAGACACTACCTAATGCTGAAATGTTGGAAGATGTTATATGCTATAA 420
OY 421 ACATATTAATAGTGATCATGTATTAACCCATCTGACTCATGTGATGAGCGCTTGAT 480
DB 421 ACATATTAATAGTGATCATGTATTAACCCATCTGACTCATGTGATGAGCGCTTGAT 480
OY 481 GTTGTTTTATACATGACCCCAATGTCCTGATGCGTTTCCCAAAATTAATGTTTAA 540
DB 481 GTTGTTTTATACATGACCCCAATGTCCTGATGCGTTTCCCAAAATTAATGTTTAA 540
OY 541 AAACGATTTGAAGCTATCCCAAAATGTAAGTACTTGAATAATCCAGCAAGTATATACCA 600
DB 541 AAACGATTTGAAGCTATCCCAAAATGTAAGTACTTGAATAATCCAGCAAGTATATACCA 600
OY 601 TGGCCTTTGACAGGCTGGCAAGCCAGCTTGGTGGTGGGACCATCTCCAAATGCGAT 660
DB 601 TGGCCTTTGACAGGCTGGCAAGCCAGCTTGGTGGTGGGACCATCTCCAAATGCGAT 660
OY 661 CTGTTCCGCGGTGATC 677
DB 661 CTGTTCCGCGGTGATC 677

RESULT 9
US-08-216-894-9
Sequence 9, Application US/08216894
Patent No. 5876734
GENERAL INFORMATION:
APPLICANT: Kirchhoff, Louis V.
APPLICANT: Otsu, Keiko
TITLE OF INVENTION: POLYPEPTIDES FOR DIAGNOSING INFECTION
TITLE OF INVENTION: WITH TRYPAOSOMA CRUZI
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.

COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/216, 894
FILING DATE: 24-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 85326/102/DRLO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1419 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1416
US-08-216-894-9

Query Match 60.5%; Score 663.4; DB 2; Length 1419;
Best Local Similarity 99.8%; Pred. No. 1.3e-179;
Matches 664; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ATGCCCCCTATCTAGGTTATGGAATAATAGAGCCCTTGCCACCCACTGACTTCT 60
DB 1 ATGCCCCCTATCTAGGTTATGGAATAATAGAGCCCTTGCCACCCACTGACTTCT 60
OY 61 TTGGAATATCTTGAGAAATAATATGAGAGCATTTGATAGCCGAGTAAAGTATAA 120
DB 61 TTGGAATATCTTGAGAAATAATATGAGAGCATTTGATAGCCGAGTAAAGTATAA 120
OY 121 TGGCCAAACAAAGTTGATGAGTTGGGTTGAGCTTCCCAATCTCTTATATATGAT 180
DB 121 TGGCCAAACAAAGTTGATGAGTTGGGTTGAGCTTCCCAATCTCTTATATATGAT 180
OY 181 GGTGATGTTAAATTAACAGCTATAGCCATCATAGCTATATAGCTGACAGCAAC 240
DB 181 GGTGATGTTAAATTAACAGCTATAGCCATCATAGCTATATAGCTGACAGCAAC 240
OY 241 ATGTTGGGTGGTTGTCACAAAGAGCGTGAGAGATTTCAATGCTTGAAGAGCGTTTG 300
DB 241 ATGTTGGGTGGTTGTCACAAAGAGCGTGAGAGATTTCAATGCTTGAAGAGCGTTTG 300
OY 301 GATATTAGATACGTTGTTGCAAAATTCGATATAGTAAAGACTTTGAAACTCTCAAGTT 360
DB 301 GATATTAGATACGTTGTTGCAAAATTCGATATAGTAAAGACTTTGAAACTCTCAAGTT 360
OY 361 GATTTTCTTAGACACTACCTAATGCTGAAATGTTGGAAGATGTTATATGCTATAA 420
DB 361 GATTTTCTTAGACACTACCTAATGCTGAAATGTTGGAAGATGTTATATGCTATAA 420
OY 421 ACATATTAATAGTGATCATGTATTAACCCATCTGACTCATGTGATGAGCGCTTGAT 480
DB 421 ACATATTAATAGTGATCATGTATTAACCCATCTGACTCATGTGATGAGCGCTTGAT 480
OY 481 GTTGTTTTATACATGACCCCAATGTCCTGATGCGTTTCCCAAAATTAATGTTTAA 540
DB 481 GTTGTTTTATACATGACCCCAATGTCCTGATGCGTTTCCCAAAATTAATGTTTAA 540
OY 541 AAACGATTTGAAGCTATCCCAAAATGTAAGTACTTGAATAATCCAGCAAGTATATACCA 600
DB 541 AAACGATTTGAAGCTATCCCAAAATGTAAGTACTTGAATAATCCAGCAAGTATATACCA 600

Db 541 AACGATATGAAGCTATCCACAAATGATTAAGTACTGAAATCCAGCAAGTATATAGCA 600

QY 601 TGCCCTTTGACAGGCTGGCAGACGATTGGTGGCGACCAATCCCAAAATCGAT 660

Db 601 TGCCCTTTGACAGGCTGGCAGACGATTGGTGGCGACCAATCCCAAAATCGAT 660

QY 661 CTGCT 665

Db 661 CTGAT 665

RESULT 10

US-08-974-180-19

Sequence 19, Application US/08974180

Patent No. 6025194

GENERAL INFORMATION:

APPLICANT: Funk, Walter

TITLE OF INVENTION: Methods for Modulating and Identifying

TITLE OF INVENTION: Cellular Senescence

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genon Corporation

STREET: 230 Constitution Drive

CITY: Menlo Park

STATE: California

COUNTRY: USA

ZIP: 94025

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/974,180

FILING DATE: 19-NOV-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Kaster, Kevin R.

REGISTRATION NUMBER: 32,704

REFERENCE/DOCKET NUMBER: 206

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 473-7779

TELEFAX: (650) 473-8654

INFORMATION FOR SEQ. ID NO: 19:

SEQUENCE CHARACTERISTICS:

LENGTH: 1920 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA

FEATURE:

NAME/KEY: CDS

LOCATION: 258..1868

OTHER INFORMATION: /product= "recombinant GS-GC6

OTHER INFORMATION: fusion protein"

US-08-974-180-19

Query Match 60.5%; Score 663.4; DB 3; Length 1920;

Best Local Similarity 99.8%; Pred. No. 1.6e-179;

Matches 664; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 181 GGTGATGTTAAATTAACACAGCTATGGCCATCATAGCTTATATAGCTGACAGACAC 240

Db 438 GGTGATGTTAAATTAACACAGCTATGGCCATCATAGCTTATATAGCTGACAGACAC 497

QY 241 ATGTTGGGTGTTGTCACAAAGAGCGTCAGAGATTTCATGCTGGAAGAGCGGTTTG 300

Db 438 ATGTTGGGTGTTGTCACAAAGAGCGTCAGAGATTTCATGCTGGAAGAGCGGTTTG 557

QY 301 GATATTATACAGGCTGTTTCGAGAAATTCGATATAGTAAAGACTTGAAGCTCAGAACTT 360

Db 558 GATATTATACAGGCTGTTTCGAGAAATTCGATATAGTAAAGACTTGAAGCTCAGAACTT 617

QY 361 GATTTCTTACAGACCTGCTGAATGCTGAAATGCTGGAAGATCGTTTGTGTCATAA 420

Db 618 GATTTCTTACAGACCTGCTGAATGCTGGAAGATCGTTTGTGTCATAA 677

QY 421 ACATATTTAAATGCTGATCATGTAAACCCATCCTGATCATGTTGTATGACGCTTGAT 480

Db 678 ACATATTTAAATGCTGATCATGTAAACCCATCCTGATCATGTTGTATGACGCTTGAT 737

QY 481 GTTGTATATACATGACCCATGCTGCTGATGCTGCTCCAAATTAAGTTTAA 540

Db 738 GTTGTATATACATGACCCATGCTGCTGATGCTGCTCCAAATTAAGTTTAA 797

QY 541 AACGATATGAAGCTATCCACAAATGATTAAGTACTGAAATCCAGCAAGTATATAGCA 600

Db 798 AACGATATGAAGCTATCCACAAATGATTAAGTACTGAAATCCAGCAAGTATATAGCA 857

QY 601 TGCCCTTTGACAGGCTGGCAGACGATTGGTGGCGACCAATCCCAAAATCGAT 660

Db 858 TGCCCTTTGACAGGCTGGCAGACGATTGGTGGCGACCAATCCCAAAATCGAT 917

QY 661 CTGCT 665

Db 918 CTGAT 922

RESULT 11

US-08-305-764C-55

Sequence 55, Application US/08305764C

Patent No. 5856090

GENERAL INFORMATION:

APPLICANT: Epstein, David M.

TITLE OF INVENTION: DNA METHYLASE LINKING REACTION

NUMBER OF SEQUENCES: 71

CORRESPONDENCE ADDRESS:

ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE

STREET: 10550 No. 5856090th Torrey Pines Road

CITY: La Jolla

STATE: California

COUNTRY: USA

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/305,764C

FILING DATE: 09-SEP-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Fitting, Thomas

REGISTRATION NUMBER: 34,163

REFERENCE/DOCKET NUMBER: TSRI 440.0

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 784-2937

TELEFAX: (619) 784-9399

INFORMATION FOR SEQ. ID NO: 55:

SEQUENCE CHARACTERISTICS:

LENGTH: 2095 base pairs

TYPE: nucleic acid

STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 9..1952
US-08-305-764C-55

Query Match 60.5%; Score 662.8; DB 2; Length 2095;
Best Local Similarity 99.0%; Pred. No. 2.4e-179;
Matches 667; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATGTCCCTATCTAGTATTGGAATAATTAAGGCGCTTGACACCCCTGACTCTT 60
DB 9 ATGTCCCTATCTAGTATTGGAATAATTAAGGCGCTTGACACCCCTGACTCTT 68
QY 61 TTGGAATATCTTGAGAAAAATATGAGAGCAATTTGTATGAGCGGATGAGGTATAA 120
DB 69 TTGGAATATCTTGAGAAAAATATGAGAGCAATTTGTATGAGCGGATGAGGTATAA 128
QY 121 TGGCGAACAACAAAGTTGATTTGGCTTTGGAGTTCCCAATCTCTTATTATTGAT 180
DB 129 TGGCGAACAACAAAGTTGATTTGGCTTTGGAGTTCCCAATCTCTTATTATTGAT 188
QY 181 GGTGATGTTAAATTAACAGAGCTATAGGCAATCATAGCTATATAGCTGACACACAC 240
DB 189 GGTGATGTTAAATTAACAGAGCTATAGGCAATCATAGCTATATAGCTGACACACAC 248
QY 241 ATGTGGGTGTTGTCACAAAGAGCGTGAGATTTCAATGCTTGAAGAGCGGTTTG 300
DB 249 ATGTGGGTGTTGTCACAAAGAGCGTGAGATTTCAATGCTTGAAGAGCGGTTTG 308
QY 301 GATATTAGTACGCTTTTCGAGATTCATATAGTAAAGCTTTGAACCTGCAAGTT 360
DB 309 GATATTAGTACGCTTTTCGAGATTCATATAGTAAAGCTTTGAACCTGCAAGTT 368
QY 361 GATTTCTTAGCAAGCTTACCTGAATGCTGAATAATGTTGAAGATTTGATGCTATAA 420
DB 369 GATTTCTTAGCAAGCTTACCTGAATGCTGAATAATGTTGAAGATTTGATGCTATAA 428
QY 421 ACATATTTAAATGATGATCATGTAAACCATCTGACTTCAATGTTGATGACGCTTGTAT 480
DB 429 ACATATTTAAATGATGATCATGTAAACCATCTGACTTCAATGTTGATGACGCTTGTAT 488
QY 481 GTTGTTTTATACATGAGACCAATGTGCTGATGCTTCCCAAAATTAAGTTTAA 540
DB 489 GTTGTTTTATACATGAGACCAATGTGCTGATGCTTCCCAAAATTAAGTTTAA 548
QY 541 AAACGATTTGAAGCTATCCCAAAATGATTAAGTACTTGAATCCAGCAAGTATATAGCA 600
DB 549 AAACGATTTGAAGCTATCCCAAAATGATTAAGTACTTGAATCCAGCAAGTATATAGCA 608
QY 601 TGGCCTTTGACAGGCTGAGCAAGCAAGTGTGCTGAGGACCATCTCCAAATGAGAT 660
DB 609 TGGCCTTTGACAGGCTGAGCAAGCAAGTGTGCTGAGGACCATCTCCAAATGAGAT 668
QY 661 CTGGTCCGCTGG 674
DB 669 ATCGAAGTCTGG 682

RESULT 12
US-08-216-894-5
Sequence 5, Application US/08216894
Patent No. 5876734
GENERAL INFORMATION:
APPLICANT: Kirchoff, Louis V.
APPLICANT: Otsu, Keiko
TITLE OF INVENTION: POLYPEPTIDES FOR DIAGNOSING INFECTION
WITH TRYPAOSOMA CRUZI

NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/216,894
FILING DATE: 24-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 85326/102/DRLO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 924 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..921
US-08-216-894-5

Query Match 60.3%; Score 661; DB 2; Length 924;
Best Local Similarity 100.0%; Pred. No. 5.3e-179;
Matches 661; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTCCCTATCTAGTATTGGAATAATTAAGGCGCTTGACACCCCTGACTCTT 60
DB 1 ATGTCCCTATCTAGTATTGGAATAATTAAGGCGCTTGACACCCCTGACTCTT 60
QY 61 TTGGAATATCTTGAGAAAAATATGAGAGCAATTTGTATGAGCGGATGAGGTATAA 120
DB 61 TTGGAATATCTTGAGAAAAATATGAGAGCAATTTGTATGAGCGGATGAGGTATAA 120
QY 121 TGGCGAACAACAAAGTTGATTTGGCTTTGGAGTTCCCAATCTCTTATTATTGAT 180
DB 121 TGGCGAACAACAAAGTTGATTTGGCTTTGGAGTTCCCAATCTCTTATTATTGAT 180
QY 181 GGTGATGTTAAATTAACAGAGCTATAGGCAATCATAGCTATATAGCTGACACACAC 240
DB 181 GGTGATGTTAAATTAACAGAGCTATAGGCAATCATAGCTATATAGCTGACACACAC 240
QY 241 ATGTGGGTGTTGTCACAAAGAGCGTGAGATTTCAATGCTTGAAGAGCGGTTTG 300
DB 241 ATGTGGGTGTTGTCACAAAGAGCGTGAGATTTCAATGCTTGAAGAGCGGTTTG 300
QY 301 GATATTAGTACGCTTTTCGAGATTCATATAGTAAAGCTTTGAACCTGCAAGTT 360
DB 301 GATATTAGTACGCTTTTCGAGATTCATATAGTAAAGCTTTGAACCTGCAAGTT 360
QY 361 GATTTCTTAGCAAGCTTACCTGAATGCTGAATAATGTTGAAGATGTTTATGCTATAA 420
DB 361 GATTTCTTAGCAAGCTTACCTGAATGCTGAATAATGTTGAAGATGTTTATGCTATAA 420
QY 421 ACATATTTAAATGATGATCATGTAAACCATCTGACTTCAATGTTGATGACGCTTGTAT 480
DB 421 ACATATTTAAATGATGATCATGTAAACCATCTGACTTCAATGTTGATGACGCTTGTAT 480
QY 481 GTTGTTTTATACATGAGACCAATGTGCTGATGCTTCCCAAAATTAAGTTTAA 540

|||||
Db 481 GTGTGTTATACATGAGCCCAATGTCCTGATGGCTCCCAAAATGTGTTTAA 540
Qy 541 AACGTTATGAACTATCCCAAAATGATAGTACTTGAATCCAGCAATATACCA 600
Db 541 AACGTTATGAACTATCCCAAAATGATAGTACTTGAATCCAGCAATATACCA 600
Qy 601 TGGCCTTTGCAAGGGCTGGCAAGCCAGTGTGTGTGGCGACCATCTCCAAATCGGAT 660
Db 601 TGGCCTTTGCAAGGGCTGGCAAGCCAGTGTGTGTGGCGACCATCTCCAAATCGGAT 660
Qy 661 C 661
Db 661 C 661

RESULT 13
US-08-895-707-1
; Sequence 1, Application US/08895707
; Patent No. 6077700
GENERAL INFORMATION:
APPLICANT: (Pharmacia & Upjohn, Co.)
APPLICANT: alternatively, for U.S. Filing:
APPLICANT: Hollingsworth, Robert A.
APPLICANT: Sharma, Satish K.
APPLICANT: Rank, Kenneth B.
APPLICANT: Evans, David B.
TITLE OF INVENTION: Special Constructs and Complexes of
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESS: Pharmacia & Upjohn Company
STREET: 301 Henrietta Street
CITY: Kalamazoo
STATE: MI
COUNTRY: USA
ZIP: 49001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Rel. #1.0, Ver. #1.25/WordPerfect 5.2+
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/895,707
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Woollon, Thomas A.
REGISTRATION NUMBER: 35,004
REFERENCE/DOCKET NUMBER: 6054
TELECOMMUNICATION INFORMATION:
TELEPHONE: 616-833-7914
TELEFAX: 616-833-8897
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1951 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
US-08-895-707-1

Query Match 60.3%; Score 660.4; DB 3; Length 1951;
Best Local Similarity 99.8%; Pred. No. 1.1e-178;
Matches 661; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGTCCCATATCTAGTATTGGAATAATTAAGGCGCTTGTCACACCCACTGACTTCT 60
Db 1 ATGTCCCATATCTAGTATTGGAATAATTAAGGCGCTTGTCACACCCACTGACTTCT 60

Qy 61 TTGGAATATCTTGAAGAAAATATGAAGACATTGTGTATGAGCCGATGATGATAA 120
Db 61 TTGGAATATCTTGAAGAAAATATGAAGACATTGTGTATGAGCCGATGATGATAA 120
Qy 121 TGGCGAACAACAAAGTTGAAATGGGTTTGAGATTTCCCAATCTTCTATTATATGAT 180
Db 121 TGGCGAACAACAAAGTTGAAATGGGTTTGAGATTTCCCAATCTTCTATTATATGAT 180
Qy 181 GGTATGTTAATTTATACACATCTATGGCCATTCATAGCTTTATATAGCGACAACAC 240
Db 181 GGTATGTTAATTTATACACATCTATGGCCATTCATAGCTTTATATAGCGACAACAC 240
Qy 241 ATGTTGGGTGGTGTGCAAGAGCGTGCAGAGATTTCAATGCTTGAAGAGCGGTTTGG 300
Db 241 ATGTTGGGTGGTGTGCAAGAGCGTGCAGAGATTTCAATGCTTGAAGAGCGGTTTGG 300
Qy 301 GATATTTAGATACGCGTGTTCAGAAATTCATATAGTAAGACTTGAAGACTCTCAAGCT 360
Db 301 GATATTTAGATACGCGTGTTCAGAAATTCATATAGTAAGACTTGAAGACTCTCAAGCT 360
Qy 361 GATTTCTTACAGCTACCTGGAATGCTGAAATGCTGAAAGTCTTATGTCATATA 420
Db 361 GATTTCTTACAGCTACCTGGAATGCTGAAATGCTGAAAGTCTTATGTCATATA 420
Qy 421 ACATATTTAATGATGATCATGTATACCCATCTGACTTCATGTTGTATGAGCGCTTGTAT 480
Db 421 ACATATTTAATGATGATCATGTATACCCATCTGACTTCATGTTGTATGAGCGCTTGTAT 480
Qy 481 GTGTGTTATATACATGAGCCCAATGCTGCTGATGCTGCCCAAAATTTGTTTAA 540
Db 481 GTGTGTTATATACATGAGCCCAATGCTGCTGATGCTGCCCAAAATTTGTTTAA 540
Qy 541 AACGTTATGAACTATCCCAAAATGATAGTACTTGAATCCAGCAATATACCA 600
Db 541 AACGTTATGAACTATCCCAAAATGATAGTACTTGAATCCAGCAATATACCA 600
Qy 601 TGGCCTTTGCAAGGGCTGGCAAGCCAGTGTGTGTGGCGACCATCTCCAAATCGGAT 660
Db 601 TGGCCTTTGCAAGGGCTGGCAAGCCAGTGTGTGTGGCGACCATCTCCAAATCGGAT 660
Qy 661 CT 662
Db 661 CT 662

RESULT 14
US-08-470-837-25
; Sequence 25, Application US/08470837
; Patent No. 5800811
GENERAL INFORMATION:
APPLICANT: Nimol, Marcel E.
APPLICANT: Hall, Frederick L.
APPLICANT: Tuan, Tai-Lan
APPLICANT: Wu, Lingtao
APPLICANT: Cheung, David T.
TITLE OF INVENTION: Transforming Growth Factor B Fusion
TITLE OF INVENTION: and
TITLE OF INVENTION: Their Use in Wound Healing
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESS: Merchant & Gould
STREET: 1150 Santa Monica Boulevard, Suite 400
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90025-3395
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:


```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match	59.7%	Score 654	DB 1	length 654
Best Local Similarity	100.0%	Pred. No.	4.4e-177	
Matches 654	Conservative 0	Mismatches 0	Indels 0	Gaps 0

QY	1	ATGTCGCCCTTACACAGGTTATTTGAAATTAAGGCGCTTGGACACCCACTCGACCTT	60
Db	1	ATGTCGCCCTTACACAGGTTATTTGAAATTAAGGCGCTTGGACACCCACTCGACCTT	60
QY	61	TTGGAATATCTTGAGAAAAATATGAGAGCAATTTGATAGAGCGCGATGAGGTATAA	120
Db	61	TTGGAATATCTTGAGAAAAATATGAGAGCAATTTGATAGAGCGCGATGAGGTATAA	120
QY	121	TGGGGAACAAAGATTGAAATTTGGGTTTGGAGTTTCCAAATCTTCTTATATATGAT	180
Db	121	TGGGGAACAAAGATTGAAATTTGGGTTTGGAGTTTCCAAATCTTCTTATATATGAT	180
QY	181	GGTGATGTTAAATTTAAGACAGCTCATGTGCGCATCATATATAGCTGACAAGCACAC	240
Db	181	GGTGATGTTAAATTTAAGACAGCTCATGTGCGCATCATATATAGCTGACAAGCACAC	240
QY	241	ATGTTGGGTGGTGTCCAAAAAGCGTGACAGATTTTCAATGCTTGAAGAGCGGTTTG	300
Db	241	ATGTTGGGTGGTGTCCAAAAAGCGTGACAGATTTTCAATGCTTGAAGAGCGGTTTG	300
QY	301	GATATTGATATCGGTGTTTGAGAAATGCAATATGTATGAACACTTGAACATCTCAAGTT	360
Db	301	GATATTGATATCGGTGTTTGAGAAATGCAATATGTATGAACACTTGAACATCTCAAGTT	360
QY	361	GATTTTCTTACAGCGTACCTGGAATGCTGAAATGTTTCAAGATCGTTTATGTCATAAA	420
Db	361	GATTTTCTTACAGCGTACCTGGAATGCTGAAATGTTTCAAGATCGTTTATGTCATAAA	420
QY	421	ACATATTTAAATGCTGATCATGTATACCCATCTGCATCTCATGTTGATATGACGCTTGAT	480
Db	421	ACATATTTAAATGCTGATCATGTATACCCATCTGCATCTCATGTTGATATGACGCTTGAT	480
QY	481	GTTGTTTTATATCATGAGCCCAATGCTGCTGATTCGTTCCCAAAATTAAGTTGTTTAA	540
Db	481	GTTGTTTTATATCATGAGCCCAATGCTGCTGATTCGTTCCCAAAATTAAGTTGTTTAA	540
QY	541	AAACGATATGAAGCTATCCCAAAATTTGATTAAGTACTTGAATTCACAGCATATATAGCA	600
Db	541	AAACGATATGAAGCTATCCCAAAATTTGATTAAGTACTTGAATTCACAGCATATATAGCA	600
QY	601	TGGCGTTTGCAGGGCTGGCAAGCCAGCTTTGGTGGCGACCATCTCTCAAAA 654	
Db	601	TGGCGTTTGCAGGGCTGGCAAGCCAGCTTTGGTGGCGACCATCTCTCTCAAAA 654	

RESULT 15
US-08-102-757-10
; Sequence 10, Application US/08102757

```

1  GENERAL INFORMATION: Masahiro
2  APPLICANT: Tajima, Masahiro
3  APPLICANT: Ohnuma, Mamami
4  APPLICANT: Ierner, Ethan A.
5  TITLE OF INVENTION: Modified Maxadilan Protein, Its
6  TITLE OF INVENTION: Preparation and Use, and DNA encoding the Protein
7  NUMBER OF SEQUENCES: 25
8  CORRESPONDENCE ADDRESS:

```

STREET: 1300 I Street, N.W., Suite 700
CITY: Washington

COMPUTER READABLE FORM:

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

```

; FILING DATE: 01-JAN-1993
 ; CLASSIFICATION: 424
 ; ATTORNEY/AGENT INFORMATION:

REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 05136.0003-000000
TELECOMMUNICATION INFORMATION:

```

; ELEMENT: 202-408-4400
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:

```

```

; MOLECULE TYPE: DNA (genomic)
US-08-102-757-10

```

Query Match	59.28;	Score 649;	DB 1;	Length 897;
Best Local Similarly	98.58;	Pred. No. 1.4e-175;		
Matches 655; Conservative	0;	Mismatches 10;	Indels 0;	Gaps 0;

Qy	1	ATGCCCCATACAGGTTATGGAAAATTAAAGGCCCTTGACAAACCACCTCACTCTT	60
Db	1	ATGTCCCTATACAGGTTATGGAAAATTAAAGGCCCTTGACAAACCACCTCACTCTT	60
Qy	61	TTGGATATCTTGAAGAAAATATGAAGACGATTTGTATAGCCGCGATGAAGTGATAAA	120
Db	61	TTGGATATCTTGAAGAAAATATGAAGACGATTTGTATAGCCGCGATGAAGTGATAAA	120
Qy	121	TGGCGAAACAAGAGTTGAAATTTGGGTTTGAGATTTCCCAATCTTCCCTAATTATATGAT	180
Db	121	TGGCGAAACAAGAGTTGAAATTTGGGTTTGAGATTTCCCAATCTTCCCTAATTATATGAT	180
Qy	181	GGTATCTTAAATTAACCAACATCTATGGCCATCATACGTTATATAGCTGACACACAAAC	240
Db	181	GGTATCTTAAATTAACCAACATCTATGGCCATCATACGTTATATAGCTGACACACAAAC	240
Qy	241	ATGTTGGGTGGTTGCCCCAAAGACGCGTCACAGATTTCAATGCTTGAAGGAGCGGTTTG	300
Db	241	ATGTTGGGTGGTTGCCCCAAAGACGCGTCACAGATTTCAATGCTTGAAGGAGCGGTTTG	300
Qy	301	GATATTACATACGGTGTTCGAGAAATTCATATGTAAAGACTTTGAAACTTCCAAAGTT	360
Db	301	GATATTACATACGGTGTTCGAGAAATTCATATGTAAAGACTTTGAAACTTCCAAAGTT	360

```

Db 301 GATATTAGATACGGTGTGTTGAGAAATGCATATAGTAAAGACTTTGAAACTCTCAAGTT 360
QY 361 GATTTTCTTAGCAAGCTACCTGAATGCTGAATAATGTTGCAAGATGTTATGTCATPAA 420
    |||
Db 361 GATTTTCTTAGCAAGCTACCTGAATGCTGAATAATGTTGCAAGATGTTATGTCATPAA 420
QY 421 ACATATTAAATGCTGATCATGTAAACCATCTGACTTCATGTTGATGACGCTTTGAT 480
    |||
Db 421 ACATATTAAATGCTGATCATGTAAACCATCTGACTTCATGTTGATGACGCTTTGAT 480
QY 481 GTTGTTTTATACATGAGACCAATGTCCTGGATGCGCTGCCAAATTTAGTTTAA 540
    |||
Db 481 GTTGTTTTATACATGAGACCAATGTCCTGGATGCGCTGCCAAATTTAGTTTAA 540
QY 541 AAACGTATTGAAGCTATCCACAAATTTGATTAAGTACTTGAATCCAGCAAGTATATAGCA 600
    |||
Db 541 AAACGTATTGAAGCTATCCACAAATTTGATTAAGTACTTGAATCCAGCAAGTATATAGCA 600
QY 601 TGGCCTTTGCAAGGCTGGCAAGCCAGCTTTGTTGTTGGCGACCATCTCCAAATGCGAT 660
    |||
Db 601 TGGCCTTTGCAAGGCTGGCAAGCCAGCTTTGTTGTTGGCGACCATCTCCAAATGCGAT 660
QY 661 CTGGT 665
    |||
Db 661 CTGAT 665

```

Search completed: March 18, 2001, 04:58:09
 Job time: 4228 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 18, 2001, 03:21:31 ; Search time 885.87 Seconds
(without alignments)
8669.677 Million cell updates/sec

Title: US-09-402-488A-1

Perfect score: 1096
Sequence: 1 ATGCCCCCTACTAGCTTA.....ATCCACTAGTTCTAGACCGC 1096

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7991742 segs, 3503743858 residues

Total number of hits satisfying chosen parameters: 15983484

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: em_estba:*
42: em_estfun:*
43: em_esthum1:*

44: em_esthum2:*
45: em_esthum3:*
46: em_esthum4:*
47: em_esthum5:*
48: em_esthum6:*
49: em_esthum7:*
50: em_esthum8:*
51: em_esthum9:*
52: em_esthum10:*
53: em_esthum11:*
54: em_esthum12:*
55: em_esthum13:*
56: em_esthum14:*
57: em_esthum15:*
58: em_esthum16:*
59: em_esthum17:*
60: em_esthum18:*
61: em_esthum19:*
62: em_esthum20:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estov1:*
68: em_estov2:*
69: em_estp11:*
70: em_estp12:*
71: em_estp13:*
72: em_estp14:*
73: em_estp15:*
74: em_estro1:*
75: em_estro2:*
76: em_estro3:*
77: em_estro4:*
78: em_estro5:*
79: em_estro6:*
80: em_estro7:*
81: em_estro8:*
82: em_estro9:*
83: em_estro10:*
84: em_estro11:*
85: em_estro12:*
86: em_estro13:*
87: gb_est41:*
88: gb_est42:*
89: gb_est43:*
90: gb_est44:*
91: gb_est45:*
92: gb_est46:*
93: gb_est47:*
94: gb_est48:*
95: gb_est49:*
96: gb_est50:*
97: gb_est51:*
98: gb_est52:*
99: gb_est53:*
100: gb_est54:*
101: gb_est55:*
102: gb_est56:*
103: gb_est57:*
104: gb_est67:*
105: gb_est68:*
106: gb_est69:*
107: gb_est70:*
108: gb_est71:*
109: gb_est72:*
110: gb_est73:*
111: gb_est74:*
112: em_esthum21:*
113: em_esthum22:*
114: em_esthum23:*
115: em_estom1:*
116: em_estom2:*

117: em_estp16:*
118: em_estp17:*
119: em_estp18:*
120: em_estp19:*
121: em_estp20:*
122: em_estp21:*
123: em_estp22:*
124: em_estp23:*
125: em_estp24:*
126: em_estp25:*
127: em_estp26:*
128: em_estp27:*
129: em_estp28:*
130: em_estp29:*
131: em_estp30:*
132: em_estp31:*
133: em_estp32:*
134: em_estp33:*
135: em_estp34:*
136: em_estp35:*
137: em_estp36:*
138: em_estp37:*
139: em_estp38:*
140: em_estp39:*
141: em_estp40:*
142: em_estp41:*
143: em_estp42:*
144: em_estp43:*
145: em_estp44:*
146: em_estp45:*
147: em_estp46:*
148: em_estp47:*
149: em_estp48:*
150: em_estp49:*
151: em_estp50:*
152: em_estp51:*
153: em_estp52:*
154: em_estp53:*
155: em_estp54:*
156: em_estp55:*
157: em_estp56:*
158: em_estp57:*
159: em_estp58:*
160: em_estp59:*
161: em_estp60:*
162: em_estp61:*
163: em_estp62:*
164: em_estp63:*
165: em_estp64:*
166: em_estp65:*
167: em_estp66:*
168: em_estp67:*
169: em_estp68:*
170: em_estp69:*
171: em_estp70:*
172: em_estp71:*
173: em_estp72:*
174: em_estp73:*
175: em_estp74:*
176: em_estp75:*
177: em_estp76:*
178: em_estp77:*
179: em_estp78:*
180: em_estp79:*
181: em_estp80:*
182: em_estp81:*
183: em_estp82:*
184: em_estp83:*
185: em_estp84:*
186: em_estp85:*
187: em_estp86:*
188: em_estp87:*
189: em_estp88:*

190: gb_gss25:*
191: gb_gss26:*
192: gb_gss27:*
193: gb_gss28:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	636	58.0	668	87	AM191462	AM191462 JAA000493
2	607.4	55.4	662	27	AI932072	AI932072 JAA000214
3	600.2	54.8	601	87	AM186543	AM186543 JAA000443
4	510	46.5	551	87	AM231225	AM231225 JAVH0009.
5	445.2	40.6	669	38	AM017133	AM017133 EST272555
6	420.6	38.4	636	38	AM017155	AM017155 EST272577
7	363	33.1	475	27	AI976995	AI976995 EST271589
8	329.4	30.1	421	89	AM329883	AM329883 JATL0240.
9	229.6	20.9	368	15	AI068236	AI068236 EST209927
10	223.2	20.4	438	14	AA999498	AA999498 MCE3400.
11	219.8	20.1	238	15	AI083054	AI083054 JAA00053.
12	219.6	20.0	828	20	AI446859	AI446859 EPM5034 E
13	127.8	11.7	550	93	AM654220	AM654220 103527 MA
14	127	11.6	624	88	AM239463	AM239463 x04008.Y
15	119	10.9	544	37	AV590629	AV590629 AV590629
16	111.6	10.2	601	37	AV594046	AV594046 AV594046
17	110	10.0	552	135	BE752999	BE752999 205854 MA
18	109.6	10.0	550	38	AV667937	AV667937 AV667937
19	108.6	9.9	582	109	BE506028	BE506028 dc18408.Y
20	107	9.8	825	12	AA792550	AA792550 v869103.T
21	105.4	9.6	671	25	AI797367	AI797367 w67c09.X
22	102.8	9.4	868	23	AI663649	AI663649 u199h08.Y
23	102.6	9.4	676	135	BE727497	BE727497 601560236
24	102.4	9.3	964	146	W34862	W34862 mc62e02.r1
25	102.2	9.3	794	16	AI128211	AI128211 gc47h02.x
26	101.2	9.2	523	93	AM652611	AM652611 100786 MA
27	101	9.2	544	111	BE682854	BE682854 181175 MA
28	100.8	9.2	651	14	AA968299	AA968299 uc71910.Y
29	100.6	9.2	574	105	BE232478	BE232478 137501 MA
30	99.6	9.1	806	23	AI663617	AI663617 u199e07.Y
31	98.8	9.0	703	2	AA137683	AA137683 mc31g06.r
32	98.4	9.0	702	2	AA110288	AA110288 mp32d07.r
33	98.4	9.0	705	107	BE383177	BE383177 601298726
34	98.4	9.0	731	107	BE386146	BE386146 601277269
35	98.4	9.0	805	135	BE783081	BE783081 601470668
36	98.2	9.0	726	3	AA199001	AA199001 mv39d05.r
37	97.6	8.9	603	8	AA498061	AA498061 v175b01.r
38	97.6	8.9	665	10	AA701663	AA701663 z143905.S
39	97.6	8.9	884	17	AI226687	AI226687 u114e02.Y
40	97.2	8.9	787	26	AI870867	AI870867 w177c10.X
41	96.8	8.8	576	28	AL134603	AL134603 DKFZP547E
42	96.4	8.8	597	107	BE392895	BE392895 601308718
43	95.2	8.7	500	108	BE486040	BE486040 173388 BA
44	95	8.7	525	95	AM823858	AM823858 u160d11.Y
45	94.8	8.6	680	137	BE912561	BE912561 601666193

ALIGNMENTS

RESULT 1
AM191462
LOCUS JAA000493.SH3 Adult SJC 7/94 Schistosoma japonicum cDNA 5', mRNA
DEFINITION sequence.
ACCESSION AM191462
VERSION AM191462.1 GI:6467090
KEYWORDS EST.
SOURCE Schistosoma japonicum.

ORGANISM Schistosoma japonicum
 Eukaryota; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea
 1 Strigeidae; Schistosomatidae; Schistosomatidae; Schistosoma.
 AUTHORS Hu, W., Brindley, P.J., and Feng, Z.
 TITLE Expressed sequence tags from adults of Schistosoma japonicum (Anhui strain) (Hu, Brindley, Feng)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Brindley, P.J.
 Molecular Parasitology Unit
 Queensland Institute of Medical Research
 300 Herston Road, Queensland 4029, Australia
 Tel: 61 7 3362 0413
 Fax: 61 7 3362 0104
 Email: paul.begimr.edu.au
 PCR Primers
 FORWARD: M13 Forward
 BACKWARD: M13 Reverse
 Seq primer: J3 Reverse
 High quality sequence stop: 547.
 Location/Qualifiers
 1. 668
 /organism="Schistosoma japonicum"
 /strain="Chinese (Anhui) strain"
 /db_xref="taxon:6182"
 /clone_lib="Adult SJC 7/94"
 /sex="Male and female"
 /tissue_type="Whole body"
 /dev_stage="Adult worms"
 /lab_host="Mouse and rabbit"
 /note="Vector: Lambda ZAP-II XR.; Site_1: EcoR I; Site_2: XhoI I; Several hundred adult Schistosoma japonicum (Anhui, P. R. China, strain), of mixed sex, were perfused from the mesenteries of experimentally infected mice and rabbits at the Queensland Institute of Medical Research, Brisbane, Australia (QIMR), and stored for several months in liquid nitrogen. Subsequently, mRNA was isolated at the QIMR from lysates of these worms by oligo dt chromatography, using a kit from Pharmacia. The mRNA was then shipped to Clontech, Palo Alto, CA, USA, who constructed a cDNA library. First strand synthesis was primed with an oligo-dT-XhoI primer and synthesized using M-MV reverse transcriptase. Second strand synthesis was accomplished with RNase H and T4 DNA polymerase. The double stranded cDNA was ligated to EcoRI linkers, digested with EcoRI and XhoI, and ligated into the phagemid vector lambda ZAP II XR. After construction of this directional library by Clontech, it was returned to the QIMR. During analysis of the library at the QIMR, we have found that a small percentage, 2% to 3%, of the clones contain inserts that appear to be highly homologous to sequences from salmonid fishes, as determined by homology comparisons using BLAST and by Southern hybridization analysis to genomic DNA from salmon (Sigma Chemical Co., St. Louis, MO) under stringent washing conditions. The remainder of the clones appear to contain S. japonicum sequences."

BASE COUNT 204 a 109 c 144 g 211 t
 ORIGIN

Query Match 58.0%; Score 636; DB 87; Length 668;
 Best Local Similarity 99.8%; Pred. No. 3.7e-172;
 Matches 647; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 8 CTATAGCTATTTGAAATTAAGGCGCTTGTCACACCGACGCTTTTGGAAAT 67
 DB 1 CTATAGCTATTTGAAATTAAGGCGCTTGTCACACCGACGCTTTTGGAAAT 60
 QY 68 ATCTTGAACAAAATATGAGAGCATTTGTATGAGCGCATGAGTGATTAATGGCGAA 127
 DB 61 ATCTTGAACAAAATATGAGAGCATTTGTATGAGCGCATGAGTGATTAATGGCGAA 120

QY 128 ACAAAGTTGTAAGTGGCTTTGGAGTTTCCCAATCTCTTATATATGATGTGATG 187
 DB 121 ACAAAGTTGTAAGTGGCTTTGGAGTTTCCCAATCTCTTATATATGATGTGATG 180
 QY 188 TTAATAATACAGCTATAGGCTATATAGCTTTATAGCTATAGCTATAGCTATAG 247
 DB 181 TTAATAATACAGCTATAGGCTATATAGCTTTATAGCTATAGCTATAGCTATAG 240
 QY 248 GTGCTTGTCCAAAAGAGCGCTGCAGAGATTTCATAGCTTTGAGATATTA 307
 DB 241 GTGCTTGTCCAAAAGAGCGCTGCAGAGATTTCATAGCTTTGAGATATTA 300
 QY 308 GATACGCTTTTCGAGAAATGATATATAGTAAGACTTTGAACCTCAAGTTGATTTTC 367
 DB 301 GATACGCTTTTCGAGAAATGATATATAGTAAGACTTTGAACCTCAAGTTGATTTTC 360
 QY 368 TTAGCAAGCTACCTGAAATGCTGAAATGCTGAAATGCTGAAATGCTGAAATGCT 427
 DB 361 TTAGCAAGCTACCTGAAATGCTGAAATGCTGAAATGCTGAAATGCTGAAATGCT 420
 QY 428 TTAATGTGATCATGTAAACCATCTGACTTCATGTGTATGACGCTTTGATGTTT 487
 DB 421 TTAATGTGATCATGTAAACCATCTGACTTCATGTGTATGACGCTTTGATGTTT 480
 QY 488 TATACATGACCCCAATGTGCTGATGCTTCCCAAAATAGTTTGTAAATAAGCTA 547
 DB 481 TATACATGACCCCAATGTGCTGATGCTTCCCAAAATAGTTTGTAAATAAGCTA 540
 QY 548 TTGAACCTTCCCAAAATGATTAAGTACTTGAATCCAGCAAGTATATAGCTGCTT 607
 DB 541 TTGAACCTTCCCAAAATGATTAAGTACTTGAATCCAGCAAGTATATAGCTGCTT 600
 QY 608 TGCAGGCTGCGCAAGCAGCTTTGGTGGCGACCATCTCCAAAT 655
 DB 601 TGCAGGCTGCGCAAG-CACGTTTGGTGGTGGCGACCATCTCCAAAT 647

RESULT 2
 A1932072 662 bp mRNA EST 20-MAR-2005
 LOCUS A1932072 Schistosoma japonicum cDNA 5', mRNA
 DEFINITION
 ACCESSION JAA000214.SH3 Adult SJC 7/94 Schistosoma japonicum
 VERSION A1932072
 KEYWORDS A1932072.1 GI:5670786
 SOURCE EST.
 ORGANISM Schistosoma japonicum.
 Schistosoma japonicum.
 Eukaryota; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea
 1 Strigeidae; Schistosomatidae; Schistosomatidae; Schistosoma.
 AUTHORS Hu, W., Brindley, P.J., and Feng, Z.
 TITLE Expressed sequence tags from adults of Schistosoma japonicum (Anhui strain) (Hu, Brindley, Feng)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Brindley, P.J.
 Molecular Parasitology Unit
 Queensland Institute of Medical Research
 300 Herston Road, Queensland 4029, Australia
 Tel: 61 7 3362 0413
 Fax: 61 7 3362 0104
 Email: paul.begimr.edu.au
 PCR Primers
 FORWARD: M13 Forward
 BACKWARD: M13 Reverse
 Insert Length: 1200 Std Error: 0.00
 Seq primer: J3 Reverse
 High quality sequence stop: 549.
 Location/Qualifiers
 1. 662
 /organism="Schistosoma japonicum"
 /strain="Chinese (Anhui) strain"
 /db_xref="taxon:6182"
 /clone_lib="Adult SJC 7/94"

FEATURES
 source

```

/sex="Male and female"
/tissue_type="Whole body"
/dev_stage="Adult worms"
/lab_host="Mouse and rabbit"
/Note="Vector: lambda ZAP-II XR; Site-1: EcoR I; Site-2:
XhoI I; Several hundred adult Schistosoma japonicum (Anhui
, P.R. China, strain), of mixed sex, were perfused from
the mesenteries of experimentally infected mice and
rabbits at the Queensland Institute of Medical Research,
Brisbane, Australia (QIMR), and stored for several months
in liquid nitrogen. Subsequently, mRNA was isolated at the
QIMR from lysates of these worms by oligo dt
chromatography, using a kit from Pharmacia. The mRNA was
then shipped to Clontech, Palo Alto, CA, USA, who
constructed a cDNA library. First strand synthesis was
primed with an oligo-dt-XhoI-primer and synthesized using
M-MLV reverse transcriptase. Second strand synthesis was
accomplished with RNase H and T4 DNA polymerase. The
double stranded cDNA was ligated to EcoRI linkers,
digested with EcoRI and XhoI, and ligated into the
phagemid vector lambda ZAP II XR. After construction of
this directional library by Clontech, it was returned to
the QIMR. During analysis of the library at the QIMR, we
have found that a small percentage, 2% to 3%, of the
clones contain inserts that appear to be highly homologous
to sequences from salmonoid fishes, as determined by
homology comparisons using BLAST and by Southern
hybridization analysis to genomic DNA from salmon (Sigma
Chemical Co., St. Louis, MO) under stringent washing
conditions. The remainder of the clones appear to contain
S. japonicum sequences."
BASE COUNT      206 a      107 c      142 g      207 t
ORIGIN

```

```

Query Match      55.4%; Score 607.4; DB 27; Length 662;
Best Local Similarity 99.5%; Pred. No. 6.3e-164;
Matches 630; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

```

```

QY 23 GGAATAATTAGGCGCTTGGCAACCCACTGACTCTTTTGGAAATATCTTGAAGAAAT 82
    |||||||
Db 1 GGAAATTAAGGCGCTTGGCAA-CCACTGACTCTTTTGGAAATATCTTGAAGAAAT 59
    |||||||
QY 83 ATGAAGAGCATTTGATGAGCGGAGTGAAGTGAATATGGCAACAAAGTTGAAT 142
    |||||||
Db 60 ATGAAGAGCATTTGATGAGCGGAGTGAAGTGAATATGGCAACAAAGTTGAAT 119
    |||||||
QY 143 TGGGTTTGGAGTTCCCATCTTCTTATATATGATGATGATGATTAATTAACACAGT 202
    |||||||
Db 120 TGGGTTTGGAGTTCCCATCTTCTTATATATGATGATGATGATTAATTAACACAGT 179
    |||||||
QY 203 CTATGGCCATCATAGCTATATAGCTGACACACACATGTTGGTGTCCAAAG 262
    |||||||
Db 180 CTATGGCCATCATAGCTATATAGCTGACACACACATGTTGGTGTCCAAAG 239
    |||||||
QY 263 AGCGTGCAGAGATTTCAATGCTTGAAGAGCGGTTTGGATTAATAGATCGGTTTGA 322
    |||||||
Db 240 AGCGTGCAGAGATTTCAATGCTTGAAGAGCGGTTTGGATTAATAGATCGGTTTGA 299
    |||||||
QY 323 GAATTCGATATAGTAAGACATTTGAACCTCAAGTTGATTTCTTGAAGTACCTG 382
    |||||||
Db 300 GAATTCGATATAGTAAGACATTTGAACCTCAAGTTGATTTCTTGAAGTACCTG 359
    |||||||
QY 383 AAATCTGATAAATGTTGAAGATGTTTATGCTATAAACAATATTTAAATGTCATCATG 442
    |||||||
Db 360 AAATCTGATAAATGTTGAAGATGTTTATGCTATAAACAATATTTAAATGTCATCATG 419
    |||||||
QY 443 TAACCATCTGACTTCATGTTGATGAGCGCTTGAATGTTGTTTATACATGACCCAA 502
    |||||||
Db 420 TAACCATCTGACTTCATGTTGATGAGCGCTTGAATGTTGTTTATACATGACCCAA 479
    |||||||
QY 503 TGTGCTGATGATGCTGCCAAATATAGTTGTTTAAACGATATGAAGCTATCCAC 562
    |||||||

```

```

Db 480 TGTGCTGATGCTGCCAAATATAGTTGTTTAAACGATATGAAGCTATCCAC 539
QY 563 AAATGATAGTACTTGAATCCAGCAAGTATATAGCATGGCTTTGCGAGGCTGGCAAG 622
    |||||||
Db 540 AAATGATAGTACTTGAATCCAGCAAGTATATAGCATGGCTTTGCGAGGCTGGCAAG 599
    |||||||
QY 623 CCAGCTTGTGTGTGGCGACCATCTCCAAAT 655
    |||||||
Db 600 -CACGTTTGTGTGTGGCGACCATCTCCAAAT 631

```

```

RESULT 3
AM186543
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

FEATURES

```

source
/organism="Schistosoma japonicum"
/strain="Chinese (Anhui) strain"
/db_xref="taxon:6182"
/clone_lib="Adult SJC 7/94"
/sex="Male and female"
/tissue_type="Whole body"
/dev_stage="Adult worms"
/Note="Vector: lambda ZAP-II XR; Site-1: EcoR I; Site-2:
XhoI I; Several hundred adult Schistosoma japonicum (Anhui
, P.R. China, strain), of mixed sex, were perfused from
the mesenteries of experimentally infected mice and
rabbits at the Queensland Institute of Medical Research,
Brisbane, Australia (QIMR), and stored for several months
in liquid nitrogen. Subsequently, mRNA was isolated at the
QIMR from lysates of these worms by oligo dt
chromatography, using a kit from Pharmacia. The mRNA was
then shipped to Clontech, Palo Alto, CA, USA, who
constructed a cDNA library. First strand synthesis was
primed with an oligo-dt-XhoI-primer and synthesized using
M-MLV reverse transcriptase. Second strand synthesis was
accomplished with RNase H and T4 DNA polymerase. The
double stranded cDNA was ligated to EcoRI linkers,
digested with EcoRI and XhoI, and ligated into the
phagemid vector lambda ZAP II XR. After construction of
this directional library by Clontech, it was returned to
the QIMR. During analysis of the library at the QIMR, we
have found that a small percentage, 2% to 3%, of the
clones contain inserts that appear to be highly homologous
to sequences from salmonoid fishes, as determined by."

```



```

OY 481 GTTGTATATACATGACCAATGTGCTG 510
Db 522 GTTGTATATACATGACCAATGTGCTG 551

RESULT 5
AM017133 669 bp mRNA EST 10-SEP-1999
LOCUS EST272555 Schistosoma mansoni male, Phil Loverde/Joe Merrick
DEFINITION Schistosoma mansoni cDNA clone SMAV95 5' end, mRNA sequence.
ACCESSION AM017133
VERSION AM017133.1 GI:5869156
KEYWORDS EST.
SOURCE Schistosoma mansoni.
ORGANISM Eukaryota; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea
REFERENCE 1 (bases 1 to 669)
AUTHORS Merrick,J.M., Osman,A., Loverde,P.T., Chandra,I., Glodek,A., Fraser
,C.M. and Lee,N.H.
TITLE Generation of a Schistosoma sp. EST (SMEST) Catalog & Schistosoma
JOURNAL Unpublished (1998)
COMMENT Contact: Norman H. Lee
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3529
Fax: 301 838 0208
Email: nhlee@ligr.org
Seq primer: M13 Reverse.
FEATURES
source location/Qualifiers
1..669
/organism="Schistosoma mansoni"
/db_xref="taxon:6183"
/clone="SMAV95"
/clone_lib="Schistosoma mansoni male, Phil Loverde/Joe
Merrick"
/sex="male"
/note="Vector: pBluescript SK; Site_1: EcoRI; Site_2: XhoI
; directionally cloned cDNAs"
BASE COUNT 216 a 114 c 139 g 199 t 1 others
ORIGIN

Query Match 40.6%; Score 445.2; DB 38; Length 669;
Best Local Similarity 81.3%; Pred. No. 3.1e-117;
Matches 516; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

OY 23 GGAATATTAAGGCGCTTGCAACCCGACCTCTTTGGAATATCTTGAAGAAAAT 82
Db 8 GGAATGTCAAAGGCTTGACACACACGCTCTTTGGACACCTTGAAAGAACTT 67

OY 83 ATGAAGACATTTGATGAGCGCGATGAGATGAATGAGCAACAAAAGTTGAAT 142
Db 68 ATGAGAACGTCGATGATGATGATGATGATGATGATGATGATGATGATGAT 127

OY 143 TGGGTTGAGTTGCCCATCTCTCTTATATATGATGATGATGATGATGATGAT 202
Db 128 TAGGCTGAGTTCCCAATCTCTCTTATATATGATGATGATGATGATGATGAT 187

OY 203 CTATGGCATCATATGATATATATGATGATGATGATGATGATGATGATGAT 262
Db 188 CTATGGCATCATATGATATATATGATGATGATGATGATGATGATGATGAT 247

OY 263 AGCGTCAGACATTTCAATGCTTGAAGAGCGGTTTGGATATGATATGATAT 322
Db 248 AACGTCGGAATATTCGATCTTGAAGAGCGGTTTGGATATGATATGATAT 307

OY 323 GAATTCATATAGTAAACACTTGAACACTGCAACAGTATGATTTCTTACAGCT 382
Db 308 GAATTCATATAGTAAACACTTGAACACTGCAACAGTATGATTTCTTACAGCT 367

```

```

OY 383 AATGCTGAATAATGTTGAGAGATCGTTATGTCATTAACATATTTAAATGATCAG 442
Db 368 GGGAGGTGAATAATGTTGAGAGATCGTTTGTCTTAACAAATATTTGAACGTAATG 427

OY 443 TAACCATCTGACCTGATGATGATGATGATGATGATGATGATGATGATGATG 502
Db 428 TAACATCTGACCTGATGATGATGATGATGATGATGATGATGATGATGATG 487

OY 503 TGTGCTGATCGCTTCCCAAAATAGTTGTTTAAACAGTATGATATCCAC 562
Db 488 AGTGCTTGAACGAGTTCCCAAAATAGTTTCTTCAAAAAGTATGATATACAC 547

OY 563 AATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 622
Db 548 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 607

OY 623 CCAGCTTTGTTGTTGCGACCATCTCCCAAAATG 657
Db 608 CTACGTTTGTGTTGCGACCATCTCCCAAAATAG 642

RESULT 6
AM017155 636 bp mRNA EST 10-SEP-1999
LOCUS EST272577 Schistosoma mansoni male, Phil Loverde/Joe Merrick
DEFINITION Schistosoma mansoni cDNA clone SMAV14 5' end, mRNA sequence.
ACCESSION AM017155
VERSION AM017155.1 GI:5869178
KEYWORDS EST.
SOURCE Schistosoma mansoni.
ORGANISM Eukaryota; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea
REFERENCE 1 (bases 1 to 636)
AUTHORS Merrick,J.M., Osman,A., Loverde,P.T., Chandra,I., Glodek,A., Fraser
,C.M. and Lee,N.H.
TITLE Generation of a Schistosoma sp. EST (SMEST) Catalog & Schistosoma
JOURNAL Unpublished (1998)
COMMENT Contact: Norman H. Lee
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3529
Fax: 301 838 0208
Email: nhlee@ligr.org
Seq primer: M13 Reverse.
FEATURES
source location/Qualifiers
1..636
/organism="Schistosoma mansoni"
/db_xref="taxon:6183"
/clone="SMAV14"
/clone_lib="Schistosoma mansoni male, Phil Loverde/Joe
Merrick"
/sex="male"
/note="Vector: pBluescript SK; Site_1: EcoRI; Site_2: XhoI
; directionally cloned cDNAs"
BASE COUNT 203 a 107 c 124 g 202 t
ORIGIN

Query Match 38.4%; Score 420.6; DB 38; Length 636;
Best Local Similarity 82.3%; Pred. No. 3.7e-110;
Matches 483; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

OY 69 TCTTGAAGAAATATGAGAGCATTTGATGAGCGCGATGAGATGATTAATGCGGAAA 128
Db 8 TCTTGGCGAAGCATATGAGAGCGCTTGTATGATGATGATGATGATGATGATGAT 67

OY 129 CAAAAGTTGAATGGGTTTGGAGTTTCCCAATCTCTCTTATATATGATGATGAT 188
Db 68 CGAAAGTTAAATTTGGCTGATTTCTTCAATCTTATATATGATGATGATGAT 127

OY 189 TAAATTAACACAGTCTATGCGCCATCATACGTTATATAGCTGACAGCACACATGTTGG 248

```


Db 128 TAAATTAACACATCTATGGCTATTTTACGTATATAGCTGACAAACATATATGTGGG 187
 Oy 249 TGGTGTCCAAAGAGCGCGAGAGATTTTCAATGCTTGAGAGCGCGTTTGGATATTAG 308
 Db 188 TGGTGTCTCAAGAGCGCGAGAAATTTCTATGCTTGAGGAGCGCATTTTATGATATTAG 247
 Oy 309 ATAGCGTGTCTTGAGAAATTCATATATAGTAAGACATTTTGAACCTCAAGATTGATTTCT 368
 Db 248 ATATGCTGTCTTCAAGAAATTCATATATAGTAAGAAATTTGAACCTCAAGATTGATTTCT 307
 Oy 369 TAGCAAGCTACTGAAATGCTGAAATGCTGCAACATGCTTTATGCTATTAACATATT 428
 Db 308 CAACCAACTCCCTGAAAGCTGAAATGCTTTGAAAGATGCTTTATCAACAAATCATATT 367
 Oy 429 AATAGTGTATCATGTAACCATCTGACATGCTTATGCTGAGCGCTGATGATTTT 488
 Db 368 GACGCGTATTAAGTATGATCATCTGACTTCATGCTGTAACGCGCTCTTGATGATGTTT 427
 Oy 489 ATACATGAGACCAATGTCCTGATGCGTCCCAAAATTAATGTTTAAACAAAGTAT 548
 Db 428 ATACATGAGACCAATGTTTGGATTAATTTTCCAAACATTAATGCTTCAACAAAGTAT 487
 Oy 549 TGAACCTATCCCAAAATTTGATTAATGCTGAAATTCAGCAAGTATATGATGCTTT 608
 Db 488 AGAAATTTTACCAACATCAAGACATGATGATTTCTAAGACATTAATGATGCTCT 547
 Oy 609 GCAGGCGTGGCAAGCGCTTTGGTGGCGGACCATCTCCCAAT 655
 Db 548 ACNAAGTTGGTCCGCACTTTTGGTGGAGATGCTCTCCCAAT 594

RESULT 7
 LOCUS A1976995 475 bp mRNA EST 27-AUG-1999
 DEFINITION EST2721589 Schistosoma mansoni male, Phil Loverde/Joie Merrick
 Schistosoma mansoni cDNA SMDA18 5' end, mRNA sequence.
 ACCESSION A1976995
 VERSION A1976995.1 GI:5790167
 KEYWORDS EST.
 SOURCE Schistosoma mansoni.
 ORGANISM Schistosoma mansoni.
 Eukaryota; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea
 : Strigeidida; Schistosomatidae; Schistosomatidae; Schistosoma.
 REFERENCE Merrick, J.M., Osman, A., Loverde, P.T., Chandra, I., Glodek, A., Fraser
 : C.M. and Lee, N.H.
 TITLE Generation of a Schistosoma sp. EST (SMEST) Catalog & Schistosoma
 Gene Index
 JOURNAL Unpublished (1998)
 COMMENT Contact: Norman H. Lee
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 3529
 Fax: 301 838 0208
 Email: nhlee@igf.org
 Seq primer: M13 Reverse.
 FEATURES
 source
 1. 475
 /db_xref="taxon:6183"
 /clone="SMDA18"
 /clone_lib="Schistosoma mansoni male, Phil Loverde/Joie
 Merrick"
 /sex="male"
 /note="Vector: pBluescript SK; Site_1: EcoRI; Site_2: XhoI
 ; directionally cloned cDNAs"
 BASE COUNT 145 a 76 c 100 g 154 t
 ORIGIN
 Query Match 33.1%; Score 363; DB 27; Length 475;
 Best Local Similarity 86.1%; Pred. No. 1.3e-93;

Matches 402; Conservative 0; Mismatches 65; Indels 0; Gaps 0;
 Oy 12 ACTAGTTATTTGAAAATTAAGGCGCTTGTGCAACCCACTGCTTTTGGATATCT 71
 Db 9 ACTGGTTATTTGAAAATTAAGGCGCTTGTGCAACCCACTGCTTTTGGATATCT 68
 Oy 72 TGAAGAAAATTAAGGCGCTTGTGCAACCCACTGCTTTTGGATATCT 131
 Db 69 TGCAGCAATTAAGGCGCTTGTGCAACCCACTGCTTTTGGATATCT 128
 Oy 132 AAGTTTGAATTTGCGGCTTGTGCAACCCACTGCTTTTGGATATCT 191
 Db 129 AAGTTTGAATTTGCGGCTTGTGCAACCCACTGCTTTTGGATATCT 188
 Oy 192 ATTTAAGCAATTTGCGGCTTGTGCAACCCACTGCTTTTGGATATCT 251
 Db 189 ATTTAAGCAATTTGCGGCTTGTGCAACCCACTGCTTTTGGATATCT 248
 Oy 252 TTGTCCAAAGAGCGCTGCAAGATTTCAATGCTTGAAGAGCGGTTTGGATATGATA 311
 Db 249 TTGTCCAAAGAGCGCTGCAAGATTTCAATGCTTGAAGAGCGGTTTGGATATGATA 308
 Oy 312 CGGTGTTGCAAGATTTGCAATGCTTGAAGAGCGGTTTGGATATGATA 371
 Db 309 TGGTGTTCAGAAATTTGCAATGCTTGAAGAGCGGTTTGGATATGATA 368
 Oy 372 CAAGTACCTGAAATTTGCAATGCTTGAAGAGCGGTTTGGATATGATA 431
 Db 369 CCAACTCCCTGCAATTTGCAATGCTTGAAGAGCGGTTTGGATATGATA 428
 Oy 432 TGGTGTTCAGAAATTTGCAATGCTTGAAGAGCGGTTTGGATATGATA 478
 Db 429 CGGTGTTGCAAGATTTGCAATGCTTGAAGAGCGGTTTGGATATGATA 475

RESULT 8
 LOCUS AM329883 421 bp mRNA EST 28-JAN-2000
 DEFINITION JAYL0240.GYL Schistosoma japonicum Lambda gt11 Express library
 Schistosoma japonicum cDNA clone JAYL0240.GY 5', mRNA sequence.
 ACCESSION AM329883
 VERSION AM329883.1 GI:6806941
 KEYWORDS EST.
 SOURCE Schistosoma japonicum.
 ORGANISM Schistosoma japonicum.
 Eukaryota; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea
 : Strigeidida; Schistosomatidae; Schistosomatidae; Schistosoma.
 REFERENCE 1 (bases 1 to 421)
 Li, Y., Wu, Z.D. and Yu, X.B.
 TITLE Expressed sequence tags from adults of Schistosoma japonicum
 (Chinese strain) (Li, Y.; Wu, Z.D.; Yu, X.B.)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Wu ZD
 Department of Parasitology
 Sun-Yat-sen University of Medical Sciences
 Box 510089, 74# Zhongshan Er Road, Guangzhou, Guangdong, P.R.China
 Tel: 86-20-87330566
 Fax: 86-20-87331679
 Email: zdwu62@163.net
 PCR PRIMERS
 FORWARD: Lambda gt11 Forward Primer
 BACKWARD: Lambda gt11 Reverse Primer
 Seq primer: Lambda gt11 Forward Primer
 High quality sequence stop: 421.
 FEATURES
 source
 1. 421
 /db_xref="taxon:6182"
 /clone="JAYL0240.GY"
 /clone_lib="Schistosoma japonicum Lambda gt11 Express
 library"
 /sex="Mix"

/note="Vector: Lambda gtl1 Sfi-Not; Site_1: EcoRI; Site_2: NotI; Several hundred adult Schistosoma japonicum(Jiangxi, P.R.China, strain), of mixed sex, were perfused from the mesenteries of experimentally infected rabbits. Double-strain cDNA synthesized with the mRNA isolated from adult worm, was inserted into the bacteriophage lambda gtl1 Sfi-Not arms between EcoRI and NotI site of the lacZ gene. The cDNA library was constructed by Chen S.Z. at Nanjing Medical University, Nanjing, Jiangsu, P.R. China. (see: Chen Shuzhen, et al. Chinese Journal of Zoonoses 1997,13(6): 23-25)"

BASE COUNT 145 a 70 c 75 g 131 t
ORIGIN

Query Match 30.1%; Score 329.4; DB 89; Length 421;
Best Local Similarity 99.7%; Pred. No. 5.8e-84;
Matches 330; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 325 ATTGCATATAGTAAGACTTTGAAACCTCAAGTTGATTTCTTACAGCAAGCTACCTGAA 384
DB 1 ATTGCATATAGTAAGACTTTGAAACCTCAAGTTGATTTCTTACAGCAAGCTACCTGAA 60
QY 385 ATGCGAAGAACTTTCGAGATCGTTTATGTCATAAACAATATTAAATGTCATCATGTA 444
DB 61 ATGCGAAGAACTTTCGAGATCGTTTATGTCATAAACAATATTAAATGTCATCATGTA 120
QY 445 ACCCATCTGACCTTATGATGACGCTCTGATGTTGTTTATACATGAGCAACCAATG 504
DB 121 ACCCATCTGACCTTATGATGACGCTCTGATGTTGTTTATACATGAGCAACCAATG 180
QY 505 TGCCCTGATCGCTTCCAAAATTAAGTTTAAAAAAGCATATGAAGCTATCCCAAA 564
DB 181 TGCCCTGATCGCTTCCAAAATTAAGTTTAAAAAAGCATATGAAGCTATCCCAAA 240
QY 565 ATTGATAGTACTTCAATCCAGCAAGTATATAGCATGCGCTTTCAGGCGTGGCAACC 624
DB 241 ATTGATAGTACTTCAATCCAGCAAGTATATAGCATGCGCTTTCAGGCGTGGCAACC 300
QY 625 ACGTTTGGTGGTGGCGACCATCTCCAAAT 655
DB 301 ACGTTTGGTGGTGGCGACCATCTCCAAAT 331

RESULT 9
A1068236/c 368 bp mRNA EST 03-AUG-1998
LOCUS A1068236 Schistosoma mansoni, Phil Loverde/Joe Merrick Schistosoma
DEFINITION mansoni cDNA clone SMC053 5' end similar to glutathione
S-transferase, mRNA sequence.
ACCESSION A1068236
VERSION A1068236 GI:3386203
KEYWORDS EST.
SOURCE Schistosoma mansoni.
ORGANISM Schistosoma mansoni.
REFERENCE Eukaryota; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea
1 (bases 1 to 368)
AUTHORS Merrick,J.M., Osman,A., Loverde,P.T., Chandra,I., Glodok,A., Fraser
C.M. and Lee,N.H.
TITLE Generation of a Schistosoma sp. EST (SMEST) Catalog & Schistosoma
Gene Index
JOURNAL Unpublished (1998)
COMMENT Contact: Merrick, J.
State University of New York at Buffalo
Department of Microbiology, School of Medicine and Biomedical
Sciences, Buffalo, NY 14214, USA
Tel: (716)-829-2158
Fax: (716)-829-2177
Email: merrick@acsu.buffalo.edu
Seq primer: M13 Reverse.
location/Qualifiers
1.368

/organism="Schistosoma mansoni"
/db_xref="taxon:6183"
/clone="SMC053"
/clone_id="Schistosoma mansoni, Phil Loverde/Joe Merrick"
/note="Vector: pBluescript SK; Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 118 a 79 c 60 g 111 t
ORIGIN

Query Match 20.9%; Score 229.6; DB 15; Length 368;
Best Local Similarity 78.1%; Pred. No. 3e-55;
Matches 289; Conservative 0; Mismatches 79; Indels 2; Gaps 1;

QY 281 TGCTGAGAGAGCGGTTTGGATATATACATCGGTGTCAGAAATTCATATAGTAAG 340
DB 368 TGCTGAGAGAGCGGTTTGGATATATACATCGGTGTCAGAAATTCATATAGTAAG 311
QY 341 ACTTGAAGACCTCTCAAAAGTATTTCTTACAGCAAGCTCAAAATGCTGAATGCTG 400
DB 310 AATATGAAGACCTCTCAAAAGTATTTCTTACAGCAAGCTCTGGAGGCTGAATAATGTCG 251
QY 401 AAGATCGTTTATGTCATAAACAATATTTAAATGGTATCATATACCAATCCCTGACTTCA 460
DB 250 AAGATCGTTTATGTCATAAACAATATTTAAATGGTATCATATACCAATCCCTGACTTCA 191
QY 461 TGTTGATAGAGCGCTTATGATGTTTATATACATGAGCAACCAATGCTGATGCTGCTCC 520
DB 190 TGTTGATAGAGCGCTTATGATGTTTATATACATGAGCAACCAATGCTGATGCTGCTCC 131
QY 521 CAAATATAGTTTGTTTAAAAAAGCATATTAAGCTATACCCCAAAATGATATGACTTGA 580
DB 130 CAAATATAGTTTGTTTAAAAAAGCATATTAAGCTATACCCCAAAATGATATGACTTGA 71
QY 581 AATCAGCAAGTATATACATGCGCTTTCAGGCGCTGGCAAGCAGCTTGGTGGCG 640
DB 70 AATCAGCAAGTATATACATGCGCTTTCAGGCGCTGGCAAGTGGATGATGATGCTGGTGGAG 11
QY 641 ACCATCTCC 650
DB 10 ATACTCTCC 1

RESULT 10
AA999498/c 438 bp mRNA EST 28-SEP-1999
LOCUS AA999498 Schistosoma mansoni, cercariae Schistosoma mansoni
DEFINITION MCB3400.M1F Schistosoma mansoni, cercariae Schistosoma mansoni
cDNA clone SMC3400 3' end similar to glutathione S-transferase,
mRNA sequence.
ACCESSION AA999498
VERSION AA999498 GI:3190053
KEYWORDS EST.
SOURCE Schistosoma mansoni.
ORGANISM Schistosoma mansoni.
REFERENCE Eukaryota; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea
1 (bases 1 to 438)
AUTHORS Santos,T.M., Johnston,D.A., Azevedo,V., Ridgers,I.L., Martinez,M.F.,
Marotta,G.B., Santos,R.L., Fonseca,S.J., Ortega,J.M., Rabeiro
,D. and Pena,S.D.J.
TITLE Analysis of the gene expression profile of Schistosoma mansoni
cercariae using the expressed sequence tag approach
JOURNAL Mol. Biochem. Parasitol. 103 (1), 79-97 (1998)
COMMENT Laboratório de Genética-Bioquímica, Departamento de Bioquímica e
Imunologia
Instituto de Ciencias Biológicas, Universidade Federal de Minas
Gerais
Avenida Antonio Carlos 6627, Belo Horizonte, MG, Brazil, 31270-010
Tel: (5531)4415611
Fax: (5531)4415409
Email: santos@mono.icb.ufmg.br


```

VERSION      AI446859.1  GI:6651649
KEYWORDS     EST
SOURCE       Echinostoma paraensei.
ORGANISM     Echinostoma paraensei.
REFERENCE    1 (bases 1 to 828)
AUTHORS      Adema,C.M., Leonard,P.M., DeJong,R.J., Day,H.L., Edwards,D.J.,
              Burell,G., Hertel,L.A. and Loker,E.S.
              Analysis of messages expressed by Echinostoma paraensei miracidia
              and sporocysts, obtained by random EST sequencing
              J. Parasitol. 86 (1), 60-65 (2000)
JOURNAL      20163500
MEDLINE
COMMENT      Contact: Loker ES
              Parasitology
              Biology, University of New Mexico
              Caspeter Hall, Albuquerque, NM 87131, USA
              Tel: 505 277 5508
              Fax: 505 277 0304
              Email: esloker@unm.edu
              Insert Length: 828   Std Error: 0.00
              Seq primer: T3 and T7.
              Location/Qualifiers
FEATURES
  source      1..828
              /organism="Echinostoma paraensei"
              /strain="New Mexico, Loker"
              /db_xref="taxon:48215"
              /clone_lib="EPMs Lambda ZAP Express"
              /sex="hermaphrodite"
              /dev_stage="miracidia and sporocysts"
              /note="Vector: Lambda ZAP Express; Site_1: Xho I; Site_2:
              EcoRI"
BASE COUNT   256 a 165 c 169 g 238 t
ORIGIN
Query Match      20 0%; Score 219.6; DB 20; Length 828;
Best Local Similarity 58.6%; Pred. No. 3e-52;
Matches 381; Conservative 0; Mismatches 269; Indels 0; Gaps 0;
QY 1 ATGTCCCTATCTAGGTATTTAGGAAATTAAGGCTTGCAACCCACGCTCTT 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 62 ATGGCTCGTGTAGATATGGCTCGCTGGTGACGATGCCAACCAATTTGCTTTT 121
QY 61 TTGAATATCTTGAAGAAAATATGAAGAGCATTTGTATGACGGGAGAGAGTATAA 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 122 CTGGATATCTCGAAGAACATACGAGAAACGATATGATGATGCGGGA 181
QY 121 TGGCAACAAAAGTTGAATTTGGGTTGGAGTTTCCCAATCTCTTATATTTGAT 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 182 TGGCAACGGAGAACACTCCATTAACCTTGATTTTCCAAACCTTCCTTATACATGAT 241
QY 181 GGTGATGTTAAATTAACAGAGCTATGCGCATCATACCTTATATAGTGACAGACAAC 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 242 GGTGATCTTTGATATCTGAGCTGCTGGCAATTAATCAGATCTCCGATAGACAAAT 301
QY 241 ATGTTGGGTGTTGCCAAAAGCGCTGACAGATTTAAATGCTTGAAGAGCGGTTTGG 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 302 CTCTGGGTGACGACCCCAAGAGACGCTCATATATACATATGCTGAGAGTGCGGTGACT 361
QY 301 GATATTATATAGGTTGTTGCAATTCATATAGTAAGCTTGAAGACTTCAAACTT 360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 362 GATTATAGATGGCAATTCGCAACAATGCTTACAGCTCTCAGATGGAAGAGAGAAAGCT 421
QY 361 GATTTTCTTACCAAGCTACCTGAAGTCTGAAGATGTTGCAAGATCTTTATGTCATAAA 420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 422 AGTGTATCGAAGAACTCCGACGCTGAAAGAGTTGCAAAAGTATTTGGAATCGAAA 481
QY 421 ACATATTTAAATGATGATCATACCATCTCTGACCTCATGTTGTATAGACGCTCTTGAT 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 482 CCAATTCCTAAAGGAGCAACCAACCTGATGATGATGATGATGATGATGATGATGAT 541
QY 481 GTTGTATTATACATGAGCAACCAATGCTGATGATGATGATGATGATGATGATGAT 540

```

```

DB 542 TGATCAAGTACGTGATGATTCMAAACATTAGATTCATTTCTTACGTGGAAGACTACATA 601
QY 541 AAACGATTTGAAGCATCTCCCACAATTTGATGATGATGATGATGATGATGATGATGAT 600
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 602 AACAGATATGAATAATTTACCCCAATCAAAAGCATATATATGATGATGATGATGATGAT 661
QY 601 TGAGCTTTGAGGCGGCGGACCGGCTTGGTGGTGGGACCATCTTC 650
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 662 TGCGCACATGATGCTGCTGTTCTGCTGATTTGGTGGGAGCAACGATCTCC 711
RESULT 13
AM654220 550 bp mRNA EST 14-JUL-2000
LOCUS     AM654220
DEFINITION 103527 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION AM654220
VERSION   AM654220.1 GI:7420046
KEYWORDS  EST.
SOURCE    cow.
ORGANISM  Bos taurus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
          Bovidae; Bovinae; Bos.
          1 (bases 1 to 550)
          Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M.,
          Bennett,G.A., Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Jaegreid
          ,W.W. and Keele,J.W.
          Design and use of four pooled tissue normalized cDNA libraries for
          EST discovery in cattle
          Unpublished (2000)
          Contact: Smith TPL
          USDA, ARS, US Meat Animal Research Center
          PO Box 166, Clay Center, NE 68933-0166, USA
          Tel: 402 762 4366
          Fax: 402 762 4390
          Email: smith@email.marc.usda.gov
          Single pass sequencing. Bases called and alt trimmed with phred
          v0.980904.e. Vector identified by cross_match with the -minscore 18
          and -mismatch 12 options.
          PCR primers
          FORWARD: AGGAACAGCTATGACCAT
          BACKWARD: GTTTCCAGTCACGACG
          Plate: 94 row: F column: 6
          Seq primer: ATTATGCTGACATATAG.
FEATURES
  source      1..550
              /organism="Bos taurus"
              /db_xref="taxon:9913"
              /clone_lib="MARC 1BOV"
              /tissue_type="pooled"
              /lab_host="DH10B"
              /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
              Library made from pooled tissue from lymph node, ovary,
              fat, hypothalamus, and pituitary."
BASE COUNT   135 a 142 c 148 g 125 t
ORIGIN
Query Match      11.7%; Score 127.8; DB 93; Length 550;
Best Local Similarity 54.2%; Pred. No. 6.9e-26;
Matches 292; Conservative 0; Mismatches 232; Indels 15; Gaps 1;
QY 1 ATGTCCCTATCTAGGTATTTAGGAAATTAAGGCTTGCAACCCACGCTCTT 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 7 ATGCCCATGATCTCGGTACTGCGGACATCCGGGCTGGCCCTGCAATCCGCTTCTC 66
QY 61 TTGAATATCTTGAAGAAAATATGAAGAGCATTTGTATGAGCGG----- 106
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 67 CTGAGATACACAGACACAACATATAGAGAGCGGACGATACGCTGATAGAGATGCTCCGAC 126
QY 107 -ATGAAGTATTAATGCGCAACAAAAGTTGAATTTGGGTTTGAAGTTTCCCATCTT 165
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 127 TATGACAGACGACGCTGCTGAATGAAAAATTCACCTGGCCCTGGACTTCCCAATCTG 186
Oy 166 CCTTATTAATGATGATGATGATTAATTAACACAGCTCTATGACCATATACGTATATA 225
    || || || || || || || || || || || || || || || || || || || || || ||
Db 187 CCTACTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 246
Oy 226 GCTGACACGACACATGTTGGTGTGTTGTCACAAAGAGCGTGCAGAGATTCATGCTT 285
    || || || || || || || || || || || || || || || || || || || || || ||
Db 247 GCTCCGACACACACACCTGCTGGGAGACAGAGAGAGATGATGCTGAGACATTTTG 306
Oy 286 GAAGAGCGGCTTTGATTAATGATAGATGAGGCTTTCAGATTCATATGAAGACTTT 345
    || || || || || || || || || || || || || || || || || || || || || ||
Db 307 GAGAACACGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 366
Oy 346 GAACTCTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 405
    || || || || || || || || || || || || || || || || || || || || || ||
Db 367 GAGAACTGACAGCTGCTTCTTGAAGAGATGCCGAAATAATCAAGCTGTTCTGACAG 426
Oy 406 CGTTTATGATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 465
    || || || || || || || || || || || || || || || || || || || || || ||
Db 427 TTTCTGCGGACGAGGCTGCTTTCGACGAGGACACCTCATGTCGATTTCTGCTT 486
Oy 466 TATGACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 524
    || || || || || || || || || || || || || || || || || || || || || ||
Db 487 TATGACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 545

RESULT 14
LOCUS AM239463 624 bp mRNA EST 13-DEC-1999
DEFINITION x040c08.y1 NCI-CGAP_Lu31 Homo sapiens cDNA clone IMAGE:2578766 5'
            similar to gb:M63509 GLUTATHIONE S-TRANSFERASE MUSCLE (HUMAN);,
            mRNA sequence.
ACCESSION AM239463
VERSION AM239463.1 GI:6571853
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 624)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
        Tumor Gene Index
        Unpublished (1997)
JOURNAL Contact: Robert Strausberg, Ph.D.
        Tel: (301) 496-1550
        Email: Robert_Strausberg@nih.gov
        Tissue Procurement: ATCC cDNA Library Preparation: Life
        Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The
        I.M.A.G.E. Consortium DNA Sequencing by: Washington University
        Genome Sequencing Center
        Clone distribution: NCI-CGAP clone distribution information can be
        found through the I.M.A.G.E. Consortium/LLNL at:
        www-bio.llnl.gov/bdirp/image/image.html
        Seq primer: -40RP from Gibco
        High quality sequence stop: 435.
FEATURES
        Location/Qualifiers
            1..624
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:2578766"
                /clone_1fb="NCI-CGAP_Lu31"
                /sex="male"
                /dev_stage="fetal, 14 wk post-conception"
                /lab_host="DH10B"
                /note="Organ: lung, cell line: Vector: pCMV-SPORT6;
                Site_1: EcoRV; Site_2: NotI; Cloned unidirectionally, no
                5' adaptor. Primer: Oligo dt. Full-length library
                constructed by life technologies."
BASE COUNT 166 a 166 c 154 g 138 t
ORIGIN

```

```

Query Match 11.6%; Score 127; DB 88; Length 624;
Best Local Similarity 51.8%; Pred. No. 1,2e-25;
Matches 323; Conservative 0; Mismatches 285; Indels 15; Gaps 1;

Oy 1 ATGTCCCTTACTAGTATTGGAATTAAGGCTTGTGACACCACTGACTTCTT 60
    || || || || || || || || || || || || || || || || || || || || || ||
Db 2 ATGCCATGACACTGGGATCTGAAACATCCGGGGGTGCGCCATCCCTGCTC 61
Oy 61 TTGGAATATCTTACAGAAATATGAAGACATTTGTATGAGCGCGATGAAGGATAAA 120
    || || || || || || || || || || || || || || || || || || || || || ||
Db 62 CTGGAATACACAGACTACAGCTACAGAGAAAGATGACATGAGGAGGCTCTGAT 121
Oy 121 T-----GGCGAAACAAAGTTGAAATGGTTGGATTGGATTTCCCACTT 165
    || || || || || || || || || || || || || || || || || || || || || ||
Db 122 TATGACAGAGCAGCAGGCGCTGAAATGAATAATTCAGCTGGGCCCTGCACTTTCCCAATCTG 181
Oy 166 CCTTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 225
    || || || || || || || || || || || || || || || || || || || || || ||
Db 182 CCTACTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 241
Oy 226 GCTGACAGACACAACTATGTTGGTGTGTCACAAAGAGCGTGCAGAGATTTCATGCTT 285
    || || || || || || || || || || || || || || || || || || || || || ||
Db 242 GCCCGACAGCACAACCTGTGCGGGAATCAGAAAGAGCAGATTGCGGAAGCAATTTTG 301
Oy 286 GAAGAGCGGCTTTTGATTAATGATACGCTGTTTCGAGAAATTCATATGTAAGACTTT 345
    || || || || || || || || || || || || || || || || || || || || || ||
Db 302 GAGAACACGTTTATGACACACCGCTATGACACTGGCCAACTCGCTATGACACCAATTTT 361
Oy 346 GAACTCTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 405
    || || || || || || || || || || || || || || || || || || || || || ||
Db 362 GAGAACTGTAACACAGATATCTGACAGGCACTCCCTGTAATGCTGTAAGCTACTCACAG 421
Oy 406 CGTTTATGATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 465
    || || || || || || || || || || || || || || || || || || || || || ||
Db 422 TTTCTGCGGACGAGGCTGCTTTCGACGAGGACACCTCATGTCGATTTCTGCTT 481
Oy 466 TATGACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 525
    || || || || || || || || || || || || || || || || || || || || || ||
Db 482 TATGATGCTCTTGAGAAACCAAGATTTTGAAGCCGACGCTGCTGATGCTTCCCAAC 541
Oy 526 TTAGTTGTTTAAATAAAGCTATTAAGATTAATTAATTAATTAATTAATTAATTAATTA 585
    || || || || || || || || || || || || || || || || || || || || || ||
Db 542 CTGAGAGACTTCACTCCCATTTGAGGCTTGAGAGAAATCTGCTACATGAAGTGC 601
Oy 586 AGCAATATATACATGCGCTTT 608
    || || || || || || || || || || || || || || || || || || || || || ||
Db 602 AGGCGCTTCTCCCAAGACCTGT 624

RESULT 15
LOCUS AV590629 544 bp mRNA EST 29-AUG-2000
DEFINITION AV590629 Bos taurus brain fetus Bos taurus cDNA clone EIBR013603
            5', mRNA sequence.
ACCESSION AV590629
VERSION AV590629.1 GI:9701622
KEYWORDS EST.
SOURCE cow.
ORGANISM Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
            Bovidae; Bos.
REFERENCE 1 (bases 1 to 544)
AUTHORS Sugimoto,Y., Hirotsune,S., Takasuga,A., Itoh,R., Jitohzono,A. and
        Suzuki,H.
TITLE bovine cDNA sequencing
JOURNAL Unpublished (2000)
COMMENT Contact: Yoshikazu Sugimoto
        Animal Genetics Division
        Shitakawa Institute of Animal Genetics
        Odakura, Nishigo, Nishi-shitakawa, Fukushima 961-8061, Japan
        Tel: 81-248-25-5641

```

Fax: 81-248-25-5725
Email: kazusugie@ccoca.ocu.ne.jp

Single pass sequencing.

This clone was obtained from a polyA-deleted cDNA library.

FEATURES

SOURCE

Location/Qualifiers
1..544

/organism="Bos taurus"

/db_xref="taxon:9913"

/clone_1id="E1BR013G03"

/clone_1lb="Bos taurus brain fetus"

/tissue.type="brain"

/dev_stage="fetus"

/lab_host="DH10B"

/note="Vector: pZL1; site_1: SalI; site_2: NotI; Poly A

was deleted from a NotI site"

BASE COUNT 131 a 141 c 148 g 124 t

ORIGIN

Query Match

Best Local Similarity 10.9%; Score 119; DB 37; Length 544;

Matches 282; Conservative 0; Mismatches 230; Indels 15; Gaps 1;

```
QY 1 ATGCCCCCTATACAGGTATTTGAAATTAAGGCGCTTGTGCAACCCACTGACTTCTT 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 18 ATGCCCATGATCCCTGGGTACTGAGACATCCGGGCTGGCCCATGCCATCCGCTTCTC 77
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 TTGGAATATCTTGAAGAAAATATGAAGCATTTGTATGAGCGCG----- 106
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 78 CTGGAGTACACAGACACAAACTATGAGAGAGGAGGAGTACTGCTAGAGATGCTCCGAC 137
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 107 -ATGAAGTGATATAATGGCAAAACAAAGTTGAATGGGTTTGGAGTTTCCCAATCTT 165
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 138 TATGACAGAGCCAGTGGCGAATGAATAATCAAGCTGGGCTGGACTTCCCAATCTG 197
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 166 CCTTATTATTTATGATGATGTTAATTACACAGTCTATGCGCATCATAGTTTATA 225
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 198 CCTTACTTATGATGATGAGTCAACAGCTCACCCAGACAGCAGCCATCTTGGTACATC 257
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 226 GCTGACAGACACAAATGTTGGTGGTGTGCCAAAAGACGTCAGAGATTCAATGCTT 285
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 258 GCTGCGAAGCAGCAACCTGTGTGGGAGACAGAGAGAGATGATGCTGTGACATTTTG 317
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 286 GAAGAGAGCGTTTGGATTTAGATTTAGATGCGTGTGAGAAATTCATATGTAAGACTTT 345
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 318 GAGAACCAAGTTATGATGATGCGCTTGGCCATGCTAGGATCTGCTACAGCCCTGACTT 377
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 346 GAAACTCTCAAGTTGATTTCTTAGCAAGTACTAGTAATGCTGTAATGTTGGAAGAT 405
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 378 GAGAAACTGAGCGCTGTTCTTGAAGAGATCCCTGAAAAAATCAAGTGTCTCAGAG 437
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 406 CGTTTATGTCATATAAATATTTAAATGATGATGTAACCATCCGACTTCATGTTG 465
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 438 TTTTGGGGGAAAGAGCGCTTGTTCAGAGGCAAGTCACTATGTGATTTCTGCTT 497
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 466 TATGAGCGCTGTGATGTTTATATACATGAGCCCAATGTGCTGGA 512
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 498 TATGAGCGCTGTGATGATGAGCCCAATGTTGAGCCCAAGTGCCTGGA 544
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: March 18, 2001, 04:33:44
Job time: 433 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2001, 04:57:22 ; Search time 77.34 Seconds
(without alignments)
151.648 Million cell updates/sec

Title: US-09-402-488a-2
Perfect score: 1850
Sequence: 1 MSPILGYWKIKGLVQPIRL.....PKQSHNDGFEEIPEEYLIQ 343

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues
Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_36:**
1: /cgn2_2/gcgdata/geneseq/geneseq/AA1980.DAT:*
2: /cgn2_2/gcgdata/geneseq/geneseq/AA1981.DAT:*
3: /cgn2_2/gcgdata/geneseq/geneseq/AA1982.DAT:*
4: /cgn2_2/gcgdata/geneseq/geneseq/AA1983.DAT:*
5: /cgn2_2/gcgdata/geneseq/geneseq/AA1984.DAT:*
6: /cgn2_2/gcgdata/geneseq/geneseq/AA1985.DAT:*
7: /cgn2_2/gcgdata/geneseq/geneseq/AA1986.DAT:*
8: /cgn2_2/gcgdata/geneseq/geneseq/AA1987.DAT:*
9: /cgn2_2/gcgdata/geneseq/geneseq/AA1988.DAT:*
10: /cgn2_2/gcgdata/geneseq/geneseq/AA1989.DAT:*
11: /cgn2_2/gcgdata/geneseq/geneseq/AA1990.DAT:*
12: /cgn2_2/gcgdata/geneseq/geneseq/AA1991.DAT:*
13: /cgn2_2/gcgdata/geneseq/geneseq/AA1992.DAT:*
14: /cgn2_2/gcgdata/geneseq/geneseq/AA1993.DAT:*
15: /cgn2_2/gcgdata/geneseq/geneseq/AA1994.DAT:*
16: /cgn2_2/gcgdata/geneseq/geneseq/AA1995.DAT:*
17: /cgn2_2/gcgdata/geneseq/geneseq/AA1996.DAT:*
18: /cgn2_2/gcgdata/geneseq/geneseq/AA1997.DAT:*
19: /cgn2_2/gcgdata/geneseq/geneseq/AA1998.DAT:*
20: /cgn2_2/gcgdata/geneseq/geneseq/AA1999.DAT:*
21: /cgn2_2/gcgdata/geneseq/geneseq/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1850	100.0	343	20	W87636 The fusion protein
2	1236	66.8	621	21	Y52289 Human HCP phospho
3	1235	66.8	257	20	W74124 GST/GI transport r
4	1235	66.8	282	20	W74123 GST/GI transport r
5	1235	66.8	422	20	Y23648 Amino acid sequenc
6	1234.5	66.7	538	19	W47001 Glutathione-S-tran
7	1234.5	66.7	564	16	R84565 Trypanosoma cruzi
8	1234.5	66.7	643	16	R84568 Trypanosoma cruzi
9	1234	66.7	282	20	W74096 GST/GI transport r
10	1233.5	66.7	824	21	Y52288 Human HCP/GST fusi
11	1232.5	66.6	262	20	W74128 GST/GI transport r
12	1232.5	66.6	282	20	W74127 GST/GI transport r

13	1232	66.6	248	20	W74113 GST/GI transport r
14	1232	66.6	248	20	W74122 GST/GI transport r
15	1232	66.6	258	20	W74117 GST/GI transport r
16	1232	66.6	259	20	W74110 GST/GI transport r
17	1232	66.6	277	20	W74116 GST/GI transport r
18	1232	66.6	277	20	W74098 GST/GI transport r
19	1231	66.5	247	20	W74103 GST/GI transport r
20	1230.5	66.5	257	20	W74126 GST/GI transport r
21	1230	66.5	247	20	W74114 GST/GI transport r
22	1230	66.5	248	20	W74101 GST/GI transport r
23	1230	66.5	258	20	W74104 GST/GI transport r
24	1230	66.5	259	20	W74130 GST/GI transport r
25	1230	66.5	259	20	W74118 GST/GI transport r
26	1230	66.5	264	20	W74129 GST/GI transport r
27	1230	66.5	267	20	W74106 GST/GI transport r
28	1229.5	66.5	257	20	W74109 GST/GI transport r
29	1229.5	66.5	277	20	W74108 GST/GI transport r
30	1229	66.4	259	20	W74125 GST/GI transport r
31	1228	66.4	247	20	W74120 GST/GI transport r
32	1228	66.4	248	20	W74112 GST/GI transport r
33	1228	66.4	248	20	W74121 GST/GI transport r
34	1228	66.4	248	20	W74100 GST/GI transport r
35	1228	66.4	248	20	W74102 GST/GI transport r
36	1228	66.4	249	20	W74115 GST/GI transport r
37	1228	66.4	257	20	W74111 GST/GI transport r
38	1228	66.4	257	20	W74119 GST/GI transport r
39	1228	66.4	257	20	W74105 GST/GI transport r
40	1228	66.4	277	20	W74099 GST/GI transport r
41	1228	66.4	277	20	W74107 GST/GI transport r
42	1228	66.4	279	20	W74097 GST/GI transport r
43	1228	66.4	282	20	W74095 GST/GI transport r
44	1227	66.3	232	16	R72793 Glutathione-S-tran
45	1227	66.3	1196	21	Y58917 Fumonisin esterase

ALIGNMENTS

RESULT	ID	W87636 standard; Protein: 343 AA.
AC	W87636;	
DE	09-MAR-1999 (first entry)	
XX		The fusion protein GST-chymosin pro-peptide-Hirudin.
XX		Fusion protein; bovine chymosin pro-peptide; leech; hirudin;
KW		anticoagulant protein; auto-catalytically maturing zymogen;
KM		carp growth hormone; vaccine.
XX		
OS		Synthetic.
OS		Bos sp.
XX		Hirudo medicinalis.
XX		
FT	Key	Location/Qualifiers
FT	Peptide	237..278
FT	Protein	/note="bovine chymosin pro-peptide"
FT		279..344
FT		/note="leech hirudin"
FT	Cleavage-site	263..264
FT		/note="pseudochymosin cleavage site"
XX		
PN	W09849326-A1.	
XX		
PD	05-NOV-1998.	
XX		
PF	23-APR-1998;	98MO-CA00398.
XX		
PR	25-APR-1997;	97US-0044254.
XX		
XX		(SEMB-) SEMBIOSYS GENETICS INC.

XX Alcantara J, Moloney M, Van Rooijen G;
 FI WPI: 1999-059646/05.
 XX N-PSDB: V83966.
 DR
 XX
 XX
 PT Preparation of recombinant polypeptides - by expression of a fusion
 protein comprising a pro-peptide derived from an autocatalytically
 maturing zymogen and a heterologous polypeptide
 Claim 27: Fig 1: 44pp; English.

XX The present sequence represents a fusion protein comprising
 CC glutathione-S-transferase (GST)-bovine chymosin pro-peptide-leech
 CC hirudin. The chymosin pro-peptide sequence is placed upstream of the DNA
 CC sequence encoding the leech anticoagulant protein hirudin. The fusion
 CC protein was made to exemplify the invention. The specification describes
 CC a method for preparing a recombinant polypeptide in a host cell. A
 CC chimeric nucleic acid sequence encoding a fusion protein comprising
 CC a pro-peptide derived from an autocatalytically maturing zymogen linked
 CC to a protein heterologous to the pro-peptide, is introduced into the host
 CC cell. The host cells are then grown to produce the fusion protein.
 CC Altering the environment of the fusion protein allows cleavage of the
 CC pro-peptide from the fusion protein to release the recombinant
 CC polypeptide. The method can be used for the preparation of recombinant
 CC proteins can be used for delivering to a human or animal a therapeutic
 CC or nutritional polypeptide such as a vaccine, a peptide antibiotic, a
 CC cattle feed enzyme, a cytokine, a gastric lipase or a lactase.
 XX
 SQ Sequence 343 AA:

Query Match 100.0%; Score 1850; DB 20; Length 343;
 Best Local Similarity 100.0%; Pred. No. 1.7e-155;
 Matches 343; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSPILGWKIKGLVOPTRLLLEYLEEKYEHLHYERDEGDKMKRKKELGEPNLPYYTD 60
 Db 1 mspllgwkwikglvoptrlllleyleekyeehlyerdegdkwnkkkrlgldfnpnlyyid 60
 QY 61 GGVKLTQSMATIRYADKNNMGGCPKERAETISMEGAVLDIRYGVSRVSKDEFTLKV 120
 Db 61 gdvkltqsmatirytadknnm19gcpkeraetismegavldirysvrsrlayskdftelkv 120
 QY 121 DFLSKLPENMKMFEDRLCKRTYNGDHVHPDMYDALDVVLYMDPKLDAFPLVCKR 180
 Db 121 dflsklpemlkmfedrlckrtkyngdhvhpdmfmdyaldvvlmdpmlclafplvckf 180
 QY 181 KRLEALPQIDKYLKSSRYTAMPLOGMATFGGSDHPKSDLVPRGSPNSRVSSGRAET 240
 Db 181 krlaalpqidkylksskylawplgwgatfsgsdhpkpsdlvprgspnsrvdssgraet 240
 QY 241 RIPIVKGKSLRKALKKEHGLEPFLKQOYGISSKYSGFVYVYDTGESQNTCLCEGSNVC 300
 Db 241 rlpivkgsllrkalkkehglepflkqoygiskskysgfvyvldctesgnqlclegsncv 300
 QY 301 CGGNKCIISGDEKKNOCVYGECTPRKPOSHNNGDPEETIEPEYIQ 343
 Db 301 ggnkcilsgdeknocvctgeptrkpsnhdgdeetiepeeyiq 343

RESULT 2
 Y52289
 ID Y52289 standard: Protein; 621 AA.
 XX
 AC Y52289;
 XX
 DT 09-FEB-2000 (first entry)
 XX
 DE Human HCP phosphodomain/GST fusion protein.
 XX
 KW Haematopoietic cell phosphatase; HCP; cloning; expression;

KW epithelial cell; growth factor receptor; interleukin-3; IL-3;
 KW erythropoietin; EPO; negative regulator; signal transduction;
 KW modulation; activity; inhibitor; stimulation; cytokine therapy;
 KW antisense therapy; gene therapy; treatment; diagnosis; phosphodomain;
 KW catalytic domain; glutathione-S-transferase; fusion.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO954450-A1.
 XX
 PD 28-OCT-1999.
 XX
 PF 15-APR-1999; 99WO-US08228.
 XX
 PR 17-APR-1998; 98US-0082072.
 XX
 PA (ORTH) ORTHO-MCNEIL PHARM INC.
 XX
 PI Jolliffe L, Barbone F, Shanahan M, Xu D;
 DR WPI: 2000-013244/01.
 XX N-PSDB: 232917.

Assay methods for modulators of hematopoietic cell phosphatase
 Claim 3; Fig 4D; 57pp; English.

This sequence represents a human haematopoietic cell
 phosphatase (HCP) catalytic domain/glutathione-S-transferase (GST)
 fusion protein. This was cloned and expressed in E. coli and
 subsequently purified for use in a variety of activity assays.
 HCP is a 68 kD non-membrane associated protein found primarily in
 haematopoietic cells, and to some extent in epithelial cells. It
 has been shown to associate with several growth factor receptors
 of receptor activation, and hence of haematopoietic signal transduction.
 The receptors it regulates include the interleukin-3 (IL-3) and
 erythropoietin (EPO) receptors. The invention relates to a novel
 method of identifying compounds that modulate HCP activity. This
 comprises combining a modulator of HCP protein activity with HCP protein
 and HCP protein substrate, and measuring an effect of the modulator on
 the activity of the HCP protein. Inhibitors of HCP can be used to
 modulate the activity of haematopoietic receptors. HCP has been shown to
 be a negative regulator of one or more signal transduction pathways in
 haematopoietic cells, and the identification of a HCP inhibitor would
 provide a synthetic stimulator to the haematopoietic system could be
 used in conjunction with other cytokine therapy, e.g., administration of
 EPO. HCP is also associated with several growth factors found in
 haematopoietic cells and it is possible that a HCP inhibitor may also
 function as a haematopoietic cell potentiator, enhancing the effect of
 growth factors, decreasing the necessary dose of the factor. The HCP
 polynucleotides are sources of probes and primers, and can be used to
 design antisense sequences, and in gene therapy. The protein, or
 active fragments thereof, may be used to treat or diagnose disorders
 which require the modulation of the HCP activity. Small molecules that
 inhibit the action of HCP can also augment the effect of EPO, which
 meaning that a lower therapeutic dose of EPO may be utilised.

Query Match 66.8%; Score 1236; DB 21; Length 621;
 Best Local Similarity 70.2%; Pred. No. 5.5e-101;
 Matches 257; Conservative 20; Mismatches 63; Indels 26; Gaps 7;

QY 1 MSPILGWKIKGLVOPTRLLLEYLEEKYEHLHYERDEGDKMKRKKELGEPNLPYYTD 60
 Db 1 mspllgwkwikglvoptrlllleyleekyeehlyerdegdkwnkkkrlgldfnpnlyyid 60
 QY 61 GGVKLTQSMATIRYADKNNMGGCPKERAETISMEGAVLDIRYGVSRVSKDEFTLKV 120
 Db 61 gdvkltqsmatirytadknnm19gcpkeraetismegavldirysvrsrlayskdftelkv 120


```

QY 121 DFLSKLPEMLKMFEDRLCHRTYLNCDHVTNPPFMYDALDVLYMDPMCLDAPKLVCFK 180
DB 121 df1sklpemlkmedrlchkytngdhvchpdtmlydaldvlymdpmcldapklycfk 180
QY 181 KRTEAIPOIDKYLKSSKYIAMPLOQWQATFGGDHPKSDLVPRGSPN----- 228
DB 181 kritea1poidkylksskylawp1qgwqatf9g9dhpksdlvprgspefayf1ltpy 240
QY 229 -SRVDSGSAEI-TRIPLKGSLSRKALKEHGLEDF--LQKQY-GISSKYSGFVVIYD 283
DB 241 actvna---adlenrvlelnkkqesedtakagfweefes1qkqevknlhqlrlegqrpenk 297
QY 284 CTESGONLCLCEGSNVCQGG-NKCILGSD-----GEKNOCVTGEGTQPKPQSHNDGDFEET 337
DB 298 gknryknllpfhshr1l1qgrdsnlpsgdylnanylnknl1pdenaktyiasqgcleat 357
QY 338 PERYLQ 343
DB 358 vndfwg 363

RESULT 3
W74123
W74124 standard; peptide: 257 AA.
AC W74124:
XX
XX 04-MAY-1999 (first entry)
DE GST/GI transport receptor binding protein fusion sequence.
XX
XX Gastro-intestinal transport receptor; binding protein; hsi; HPT1;
KW D2H; hPEPT1; human; GI tract receptor; sucrose-isomaltase complex;
KW intestinal peptide-associated transporter; hypertension; diabetes;
KW osteoporosis; haemophilia; anaemia; cancer; migraine; angina pectoris;
KW therapeutic agent delivery; therapy; fusion protein.
XX
XX Homo sapiens.
XX
XX W09851325-A2.
XX
XX 19-NOV-1998.
XX
XX 15-MAY-1998; 98WO-US10088.
XX
XX 15-MAY-1997; 97US-0046595.
XX
XX (CYTO-) CYTOGEN CORP.
XX (ELAN-) ELAN CORP PLC.
XX
XX Alvarez VL, Belinka BA, Cagney GM, Carter JM, Lambkin IJ;
XX Omahony DJ, Patterson CA, Singleton J;
XX WPI; 1999-009568/01.
XX
XX New proteins that bind specifically to receptors in the
XX gastro-intestinal tract and related nucleic acid - chimaeras and
XX antibodies; used to deliver therapeutic or diagnostic agents to, or
XX through, the gastrointestinal tract, e.g. insulin or leuprolide
XX
XX Example: Page 224; 294pp; English.
XX
XX This sequence represents a fusion protein between
XX glutathione-S-transferase and a gastro-intestinal transport protein
XX binding peptide. The invention relates to purified proteins (I) that bind
XX specifically to at least one of the gastro-intestinal (GI) tract
XX receptors human intestinal peptide-associated transporter (HPT1),
XX hPEPT1, D2H and human sucrose-isomaltase complex (hsi). (I) provide
XX active transport of therapeutic agents through human and animal GI tissue
XX (into the blood) for in vivo delivery, particularly for treatment or
XX prevention of hypertension, diabetes, osteoporosis, haemophilia, anaemia,
XX cancer, migraine, or angina pectoris. Specifically they are used to

```

```

CC deliver insulin or leuprolide, but many other suitable therapeutic agents
CC are disclosed, including genes or inhibitory nucleic acid, imaging agents
CC and antigens. (I) may also provide targeting to the GI tract. Other uses
CC of (I) are: (i) to determine the level of specified receptors in a sample
CC (in a binding assay); and (ii) to screen for molecules that bind (I).
CC Immunogenic analogues or derivatives of (I) are used to raise antibodies
CC and in immunoassays. The antibodies are used to locate, detect and
CC measure (I), e.g. for imaging, monitoring treatment, tissue analysis
CC etc., also for peptide purification and immobilisation.
CC
SQ Sequence 257 AA:

Query Match 66.8%; Score 1235; DB 20; Length 257;
Best Local Similarity 95.5%; Pred. No. 2, 2e-101;
Matches 231; Conservative 3; Mismatches 6; Indels 2; Gaps 1;

QY 1 MSPILGYWKIKGLVOPTRLLLEYLEEKYEENHLYERDEGDKWRNKKFELGEPNLPYYID 60
DB 1 mspilgywkikglvoptrllleyleekyeehlyerdegdkwrnkkfelgfeplpyid 60
QY 61 GDVKLTOSMAIIRYIADKHNMILGSCPERAEISMLGAVIDRYGVSRATYXSKDEFELK 120
DB 61 gdwkltosma1iryadi khnmilgscperae1smlegavldrygvsr1atyskdefelk 120
QY 121 DFLSKLPEMLKMFEDRLCHRTYLNCDHVTNPPFMYDALDVLYMDPMCLDAPKLVCFK 180
DB 121 df1sklpemlkmedrlchkytngdhvchpdtmlydaldvlymdpmcldapklycfk 180
QY 181 KRTEAIPOIDKYLKSSKYIAMPLOQWQATFGGDHPKSDLVPRGSPNSRVDSGSAEIT 240
DB 181 kritea1poidkylksskylawp1qgwqatf9g9dhpksdlvprgspefayf1ltpy 240
QY 241 RI 242
DB 239 tv 240

RESULT 4
W74123
W74123 standard; peptide: 282 AA.
AC W74123:
XX
XX 04-MAY-1999 (first entry)
DE GST/GI transport receptor binding protein fusion sequence.
XX
XX Gastro-intestinal transport receptor; binding protein; hsi; HPT1;
KW D2H; hPEPT1; human; GI tract receptor; sucrose-isomaltase complex;
KW intestinal peptide-associated transporter; hypertension; diabetes;
KW osteoporosis; haemophilia; anaemia; cancer; migraine; angina pectoris;
KW therapeutic agent delivery; therapy; fusion protein.
XX
XX Homo sapiens.
XX
XX W09851325-A2.
XX
XX 19-NOV-1998.
XX
XX 15-MAY-1998; 98WO-US10088.
XX
XX 15-MAY-1997; 97US-0046595.
XX
XX (CYTO-) CYTOGEN CORP.
XX (ELAN-) ELAN CORP PLC.
XX
XX Alvarez VL, Belinka BA, Cagney GM, Carter JM, Lambkin IJ;
XX Omahony DJ, Patterson CA, Singleton J;
XX WPI; 1999-009568/01.
XX
XX New proteins that bind specifically to receptors in the

```

PT gastro-intestinal tract and related nucleic acid - chimaeras and
 PT antibodies, used to deliver therapeutic or diagnostic agents to, or
 through, the gastro-intestinal tract, e.g. insulin or leuprolide
 XX
 PS Example; Page 223; 294pp; English.

CC This sequence represents a fusion protein between
 CC glutathione-S-transferase and a gastro-intestinal transport protein
 CC binding peptide. The invention relates to purified proteins (I) that bind
 CC specifically to at least one of the gastro-intestinal (GI) tract
 CC receptors human intestinal peptide-associated transporter (HPII),
 CC hPEPT1, D2H and human sucrose-isomaltase complex (hSIC). (i) provide
 CC active transport of therapeutic agents through human and animal GI tissue
 CC (into the blood) for in vivo delivery, particularly for treatment or
 CC prevention of hypertension, diabetes, osteoporosis, haemophilia, anaemia,
 CC cancer, migraine, or angina pectoris. Specifically they are used to
 CC deliver insulin or leuprolide, but many other suitable therapeutic agents
 CC are disclosed, including genes or inhibitory nucleic acid, imaging agents
 CC and antigens. (II) may also provide targeting to the GI tract. Other uses
 CC of (I) are: (i) to determine the level of specified receptors in a sample
 CC (in a binding assay); and (ii) to screen for molecules that bind (I).
 CC Immunogenic analogues or derivatives of (I) are used to raise antibodies
 CC and in immunoassays. The antibodies are used to locate, detect and
 CC measure (I), e.g. for imaging, monitoring treatment, tissue analysis
 CC etc., also for peptide purification and immobilisation.

SO Sequence 282 AA;

Query Match 66.8%; Score 1235; DB 20; Length 282;
 Best Local Similarity 95.5%; Pred. No. 2.5e-101;
 Matches 221; Conservative 3; Mismatches 6; Indels 2; Gaps 1;

QY 1 MSPILGVMKIGLVOPRLILEEYEEHLYERDEGDMKRNKKELEGEFNNLPYYID 60
 DB 1 MSPILGYWKIKGLVQPTLILLEYLEEYEEHLYERDEGDKWNKKELGLEFNNLPYYID 60
 QY 61 GGVKLTOSMAIRYIADKNNMLGGCPKERAETSMLEGAVIDIRGVSRIRAYSKDEFTLV 120
 DB 61 GGVKLTOSMAIRYIADKNNMLGGCPKERAETSMLEGAVIDIRGVSRIRAYSKDEFTLV 120
 QY 121 DFLSKLEPMKMFEDRLCHRTYLNQDHTVHPDPMYDALDVLYMDPCLDAEPKLVCFK 180
 DB 121 DFLSKLEPMKMFEDRLCHRTYLNQDHTVHPDPMYDALDVLYMDPCLDAEPKLVCFK 180
 QY 181 KRLEAIPOIDKYLKSSKYIAMPLQGWQATFGGDPKPSDLVPGSPNSRVDSGSGRAEIT 240
 DB 181 KRLEAIPOIDKYLKSSKYIAMPLQGWQATFGGDPKPSDLVPGSPNSRVDSGSGRAEIT 240
 QY 241 RI 242
 DB 239 RV 240

RESULT 5

ID Y23648 Y23648 standard; Protein; 422 AA.

AC Y23648;

DT 07-SEP-1999 (first entry)

DE Amino acid sequence of a beta-1ipotroprolin (BLT)/GSM fusion protein.

KW Analogue; beta-1ipotroprolin; BLT; type I; type II; diabetes; retinopathy;
 KW cardiomyopathy; nephropathy; blood insulin level; blood glucose level;
 KW insulin sensitivity; fusion protein.

OS Synthetic.

OS Homo sapiens.

XX EPP26239-A2.

PD 30-JUN-1999.
 XX
 PF 21-DEC-1998; 98EP-0310498.

XX 13-OCT-1998; 98US-0103976.
 PR 23-DEC-1997; 97US-0068659.
 PR 30-MAR-1998; 98US-0079857.
 PR 21-MAY-1998; 98US-0066321.
 PR 01-JUL-1998; 98US-0091385.
 PR 05-AUG-1998; 98US-0095405.

PA (ELIL) LILLY & CO ELI.

PI Becker GW, Butler JP, Hale JE, Heath WF, Heiman ML;
 PI Schoner BE, Varshavsky AD;

DR WPI; 1999-349240/30.

PT New human protein useful for preventing and treating diabetes and
 the conditions which result from it

PS Claim 1; Page 30-31; 57pp; English.

CC The specification describes a functional analogue of beta-1ipotroprolin
 CC (BLT). The BLT analogues or their fragments may be administered to
 CC patients to prevent or treat type I or type II diabetes or the
 CC complications which arise from them such as retinopathy, cardiomyopathy
 CC and nephropathy, by decreasing blood insulin levels and therefore blood
 CC glucose levels. They may also be used to increase sensitivity to insulin
 CC in patients. The present sequence represents a BLT fusion protein.

SO Sequence 422 AA;

Query Match 66.8%; Score 1235; DB 20; Length 422;
 Best Local Similarity 80.1%; Pred. No. 4.1e-101;
 Matches 245; Conservative 6; Mismatches 27; Indels 28; Gaps 4;

QY 1 MSPILGVMKIGLVOPRLILEEYEEHLYERDEGDMKRNKKELEGEFNNLPYYID 60
 DB 1 MSPILGYWKIKGLVQPTLILLEYLEEYEEHLYERDEGDKWNKKELGLEFNNLPYYID 60
 QY 61 GGVKLTOSMAIRYIADKNNMLGGCPKERAETSMLEGAVIDIRGVSRIRAYSKDEFTLV 120
 DB 61 GGVKLTOSMAIRYIADKNNMLGGCPKERAETSMLEGAVIDIRGVSRIRAYSKDEFTLV 120
 QY 121 DFLSKLEPMKMFEDRLCHRTYLNQDHTVHPDPMYDALDVLYMDPCLDAEPKLVCFK 180
 DB 121 DFLSKLEPMKMFEDRLCHRTYLNQDHTVHPDPMYDALDVLYMDPCLDAEPKLVCFK 180
 QY 181 KRLEAIPOIDKYLKSSKYIAMPLQGWQATFGGDPKPSDLVPGSPNSRVDSGSGRAEIT 240
 DB 181 KRLEAIPOIDKYLKSSKYIAMPLQGWQATFGGDPKPSDLVPGSPNSRVDSGSGRAEIT 240
 QY 241 RIPLYKGSUR-----KALKHGLLEDFLOKOO-----YGISK-- 274
 DB 241 grel-tgqrlregdgpdpdgdagadalehsllvaakekdegprymehfrgspkdk 299
 QY 275 -YSGFV 279
 DB 300 ryggim 305

RESULT 6

ID W47001 W47001 standard; Protein; 538 AA.

AC W47001;

DT 13-AUG-1998 (first entry)

DE Glutathione-S-transferase and hTFR fusion protein 1.

XX

KM Human: telomerase reverse transcriptase; hTERT; TRT; diagnosis;
 XX prognosis; cell proliferation; cancer; ageing; ribonucleoprotein.
 OS Synthetic.
 XX Homo sapiens.
 PH Key
 FT Region 1..220 Location/Qualifiers
 FT Misc-difference /note="glutathione-S-transferase fragment"
 FT 221..226 /note="thrombin cleavage sequence"
 FT Misc-difference /note="heart muscle protein kinase recognition sequence"
 FT 232..236 /note="residues introduced by cloning"
 FT 237..538 /note="hTERT protein fragment"
 FT Region
 XX GB2317891-A.
 PD 08-APR-1998.
 XX 01-OCT-1997; 97GB-0020890.
 XX 14-AUG-1997; 97US-0915503.
 PR 01-OCT-1996; 96US-0724643.
 PR 18-APR-1997; 97US-0844419.
 PR 25-APR-1997; 97US-0846017.
 PR 06-MAY-1997; 97US-0851843.
 PR 09-MAY-1997; 97US-0854050.
 PR 14-AUG-1997; 97US-0911312.
 PR 14-AUG-1997; 97US-0912951.
 XX (GERO-) GERON CORP.
 PA (UYTE-) UNIV TECHNOLOGY CORP.
 PI Andrews WH, Cech TR, Chapman KB, Harley C, Lingner J;
 PI Morin GB, Nakamura T, Harley CB;
 DR WPI: 1998-171633/16.
 XX Pure and recombinant human Telomerase Reverse Transcriptase and its
 PT variants - are useful in the diagnosis, prognosis and treatment of
 PT cell proliferation conditions especially cancer and ageing
 XX
 PS Example 6; Page 224; 387pp; English.
 CC The present sequence represents a fusion protein from an example
 CC of the present invention which describes human telomerase reverse
 CC transcriptase (hTERT). The present invention also describes the
 CC following methods: (A) determining whether a test compound is
 CC a modulator of hTERT, by detecting the change in hTERT recombinant
 CC protein or polynucleotide, on administration of the compound;
 CC (B) preparation of recombinant telomerase by contacting a protein
 CC preparation of hTERT with a telomerase RNA component; (C) detection
 CC of the hTERT RNA or protein in a sample by binding a relevant
 CC RNA detection, amplifying the product and correlating the presence of
 CC complex or amplification product with presence of hTERT in the sample;
 CC and (D) increasing the proliferation of a vertebrate cell by increasing
 CC hTERT expression; and (E) the use of an agent that causes an increase in
 CC cell vertebrate cell proliferation to create a medicament that inhibits
 CC ageing. A protein preparation of hTERT and the polynucleotide encoding
 CC hTERT can be used in the manufacture of medicaments for inhibiting the
 CC effect of ageing or cancer. Inhibitors of telomerase activity can be
 CC used to treat conditions that are associated with high telomerase
 CC activity. A protein preparation of hTERT can also be used in the new
 CC methods.
 CC Sequence 538 AA;
 XX
 Query Match 66.7%; Score 1234.5; DB 19; Length 538;

Best Local Similarity 85.1%; Pred. No. 6, 2e-101;
 Matches 240; Conservative 6; Mismatches 13; Indels 23; Gaps 4;
 QY 1 MSPILGWKIKIGLVOPRLLEYLEEKYERDGDGKWRKKFELGEPNLPYYID 60
 DB 1 mspllgwykikgltvpttrllleyleekyeelhyeregdgkwrnkfelgiefpnlpyid 60
 QY 61 GDVYKLTQSMALIRYADKNHMLGCEPERAEISMLBGAVIDIRYVSRIAYSKDFETLV 120
 DB 61 gdvyltqsmalirlyadknhmlgscperaeismlbgaavidirysrlyayskdfetlv 120
 QY 121 DFLSKLPENMKMFEDRLCHRTYINGDHVTHPDMLYDALDVLVLYMDPMCLDAPKLYCFK 180
 DB 121 dflsklpemlkmfედrlchrttyingdhvthpdmlydaldvlympdmcldapklvcfk 180
 QY 181 KRLEAIPOIDKYLKSSKYIAMPLOGMQATPGGSDHPKPSDLVGRGSPNSHVSGRAREIT 240
 DB 181 krleaipoidkylksskyiamplogmqatpggshpksdlvgrgspnshvsgrasvgs---vt 235
 QY 241 RIP-----LYKG---KSLRKALKERHGL---EDFL 264
 DB 236 klpqgsllsllsllsclygdmeklfaglrdrjllrlvdfl 277
 RESULT 7
 R84565
 ID R84565 standard; Protein; 564 AA.
 XX
 AC R84565;
 XX
 DT 09-MAY-1996 (first entry)
 XX
 DE Trypanosoma cruzi TCR27 polypeptide, Ag15.
 XX
 KW TCR27; Chagas disease; repeat unit; diagnosis; blood screening;
 KW recombinant; fusion protein; glutathione-S-transferase.
 XX
 OS Trypanosoma cruzi.
 XX
 FH Key Location/Qualifiers
 FT Region 329..552 /label="repeat_region"
 FT /note="16 of 69 repeat units of 14 amino acids"
 XX
 PN W09525797-A1.
 PD 28-SEP-1995.
 XX
 PF 20-MAR-1995; 95WO-US03191.
 XX
 PR 24-MAR-1994; 94US-0216894.
 XX
 PA (KIRC/) KIRCHHOFF L V.
 PA (OTSU/) OTSU K.
 PI Kirchhoff LV, Otsu K;
 DR WPI: 1995-344618/44.
 DR N-PSDB: T05329.
 XX
 PT New polypeptide(s) from the TCR27 protein of Trypanosoma cruzi - as
 PT immunoassay reagent for specific diagnosis of Chagas disease, also
 PT related nucleic acid and transformed cells
 XX
 PS Disclosure; Page 30-31; 68pp; English.
 CC R84565-R84569 are polypeptides of the TCR27 protein of T. cruzi
 CC The proteins are all fusion products with glutathione-S-transferase
 CC (GST) and some contain a linker sequence. The TCR27 protein comprises
 CC a 95 amino acid (aa) N-terminal region; 69 repeats of a highly
 CC conserved 14 aa sequence and a 68 aa C-terminal region. This sequence
 CC encodes the GST sequence. The Ag15 polypeptide conty. 16 of the 69
 CC repeat units and also contains a linker sequence. The TCR27

CC polypeptides of the invention are useful for the diagnosis of Chagas
 CC disease (American Trypanosomiasis), they are capable of detecting
 CC anti-T. cruzi antibodies; or for blood screening. The TCR27 protein
 CC has epitopes to which most T. cruzi infected individuals have
 CC antibodies. The TCR27 polypeptides will not react with serum from
 CC patients with leishmaniasis, schistosomiasis, or autoimmune disease
 CC and are hence less likely to cause false positives in diagnosis.

XX Sequence 564 AA;

Query Match 66.7%; Score 1234.5; DB 16; Length 564;
 Best Local Similarity 89.3%; Pred. No. 6.6e-101;
 Matches 233; Conservative 10; Mismatches 9; Indels 9; Gaps 2;

QY 1 MSPILGYWKIKGLVPTRLLEYLEEKEEHLIERDEGDKRNKKFELGLEPMLPYIID 60
 Db 1 mspilgywkikglvptrrllleyleekeehehlyerdegdkwnkkfelglefplpyiid 60
 QY 61 GDVKTQSMALIRYIADKHNMLGGCPKERAETSMLEGAVALDIRYGSRIASKEPTELKY 120
 Db 61 gdvktqsmaliryladkhnmlggcpkeraetismlegavldirysriayskdefetlky 120
 QY 121 DFLSKPEMLKMFEDRLCHKTYLNGSDHVTNPPDMLYDALDVLYLMDPCLDAFPLKVCFK 180
 Db 121 dflskpelmkmfedrlchkytngdhvthpdmlydalvlylmpdmcldafplkvcfk 180
 QY 181 KRIEAIPOIDKYLKSSKTIAMPLOGWQATFGGCHPDKSLVPRGSPNSRYD----SSG 235
 Db 181 krieaipoidkylkssktylawplogwqatfggchpdkslvprgspnsqldgaennitns 240
 QY 236 RAEIRIPIPLYKGSLSRLAKE 256
 Db 241 krentkl-----tekvkaeke 257

RESULT 8

R84568
 ID R84568 standard; Protein: 643 AA.

XX R84568;

DT 09-MAY-1996 (first entry)

DE Trypanosoma cruzi TCR27 polypeptide, Ag15.

KW TCR27; Chagas disease; repeat unit; diagnosis; blood screening;
 KM recombinant; fusion protein; glutathione-S-transferase.

XX Trypanosoma cruzi.

Key Location/Qualifiers
 Region 329..552
 /label=repeat_region

FT /note= "16 of 69 repeat units of 14 amino acids"

XX W09525797-A1.

XX 28-SEP-1995.

XX 20-MAR-1995; 95WO-US03191.

XX 24-MAR-1994; 94US-0216894.

XX (KIRC/) KIRCHHOFF L V.

XX (OTSU/) OTSU K.

XX Kirchhoff LV, Otsu K;

XX WPI: 1995-344618/44.

XX N-PSDB: T05332.

New polypeptide(s) from the TCR27 protein of Trypanosoma cruzi - as

PT immunoassay reagent for specific diagnosis of Chagas disease, also
 PT related nucleic acid and transformed cells

XX Disclosure: Page 40-41; 68pp; English.

CC R84565-R84569 are polypeptides of the TCR27 protein of T. cruzi
 CC (GST) and some contain a linker sequence. The TCR27 protein comprises
 CC a 95 amino acid (aa) N-terminal region; 69 repeats of a highly
 CC conserved 14 aa sequence, and a 68 aa C-terminal region. This sequence
 CC encodes the GST sequence, the Ag44 polypeptide conty. 16 of the 69
 CC repeat units and also contains the amino and carboxy terminal
 CC peptides of TCR27. The TCR27 polypeptides of the invention are useful
 CC for the diagnosis of Chagas disease (American Trypanosomiasis), they
 CC are capable of detecting anti-T. cruzi antibodies; or for blood
 CC screening. The TCR27 protein has epitopes to which most T. cruzi
 CC infected individuals have antibodies. The TCR27 polypeptides will not
 CC react with serum from patients with leishmaniasis, schistosomiasis,
 CC or autoimmune disease and are hence less likely to cause false
 CC positives in diagnosis.

XX Sequence 643 AA;

Query Match 66.7%; Score 1234.5; DB 16; Length 643;
 Best Local Similarity 89.3%; Pred. No. 7.8e-101;
 Matches 233; Conservative 10; Mismatches 9; Indels 9; Gaps 2;

QY 1 MSPILGYWKIKGLVPTRLLEYLEEKEEHLIERDEGDKRNKKFELGLEPMLPYIID 60
 Db 1 mspilgywkikglvptrrllleyleekeehehlyerdegdkwnkkfelglefplpyiid 60
 QY 61 GDVKTQSMALIRYIADKHNMLGGCPKERAETSMLEGAVALDIRYGSRIASKEPTELKY 120
 Db 61 gdvktqsmaliryladkhnmlggcpkeraetismlegavldirysriayskdefetlky 120
 QY 121 DFLSKPEMLKMFEDRLCHKTYLNGSDHVTNPPDMLYDALDVLYLMDPCLDAFPLKVCFK 180
 Db 121 dflskpelmkmfedrlchkytngdhvthpdmlydalvlylmpdmcldafplkvcfk 180
 QY 181 KRIEAIPOIDKYLKSSKTIAMPLOGWQATFGGCHPDKSLVPRGSPNSRYD----SSG 235
 Db 181 krieaipoidkylkssktylawplogwqatfggchpdkslvprgspnsqldgaennitns 240
 QY 236 RAEIRIPIPLYKGSLSRLAKE 256
 Db 241 krentkl-----tekvkaeke 257

RESULT 9

W74096
 ID W74096 standard; peptide; 282 AA.

XX W74096;

DT 04-MAY-1999 (first entry)

DE GST/GI transport receptor binding protein fusion sequence.

KW Gastro-intestinal transport receptor; binding protein; hsi; HPT1;
 KW DZH; hPEPT1; human; GI tract receptor; sucrose-isomaltase complex;
 KW intestinal peptide-associated transporter; hypertension; diabetes;
 KW osteoporosis; haemophilia; anaemia; cancer; migraine; angina pectoris;
 KW therapeutic agent delivery; therapy; fusion protein.

XX Homo sapiens.

XX W09851325-A2.

XX 19-NOV-1998.

XX 15-MAY-1998; 98WO-US10088.

PR 15-MAY-1997; 97US-0046595.
 XX (CYTO-) CYTOGEN CORP.
 PA (ELAN-) ELAN CORP PLC.
 XX
 PI Alvarez VL, Belinka BA, Cagney GM, Carter JM, Lambkin LJ;
 PI Omahony DJ, Patterson CA, Singleton J;
 XX WPI: 1999-009568/01.
 DR
 XX
 PT New proteins that bind specifically to receptors in the
 PT gastro-intestinal tract and related nucleic acid - chimaeras and
 PT antibodies, used to deliver therapeutic or diagnostic agents to, or
 PT through, the gastrointestinal tract, e.g. insulin or leuprolide
 XX
 PS Example: Page 204; 294pp; English.
 CC This sequence represents a fusion protein between
 CC glutathione-S-transferase and a gastro-intestinal transport protein
 CC binding peptide. The invention relates to purified proteins (I) that bind
 CC specifically to at least one of the gastro-intestinal (GI) tract
 CC receptors human intestinal peptide-associated transporter (HPT1),
 CC hPPT1, DPH and human sucrose-isomaltase complex (HSI). (I) provide
 CC active transport of therapeutic agents through human and animal GI tissue
 CC (into the blood) for in vivo delivery, particularly for treatment or
 CC prevention of hypertension, diabetes, osteoporosis, haemophilia, anaemia,
 CC cancer, migraine, or angina pectoris. Specifically they are used to
 CC deliver insulin or leuprolide, but many other suitable therapeutic agents
 CC are disclosed. Including genes or inhibitory nucleic acid, imaging agents
 CC and antigens. (I) may also provide targeting to the GI tract. Other uses
 CC of (I) are: (i) to determine the level of specified receptors in a sample
 CC (in a binding assay); and (ii) to screen for molecules that bind (i).
 CC Immunogenic analogues or derivatives of (I) are used to raise antibodies
 CC and in immunoassays. The antibodies are used to locate, detect and
 CC measure (i), e.g. for imaging, monitoring treatment, tissue analysis
 CC etc., also for peptide purification and immobilisation.
 XX
 SQ Sequence 282 AA:
 Query Match 66.7%; Score 1234; DB 20; Length 282;
 Best Local Similarity 77.8%; Pred. No. 3e-101;
 Matches 245; Conservative 4; Mismatches 20; Indels 46; Gaps 6;
 QY 1 MSPILGWMKIKGLVOTRLLLEKYEENHYERDEGDKMKRKKFELFEPNLPYYID 60
 DB 1 MSPILGWMKIKGLVOTRLLLEKYEENHYERDEGDKMKRKKFELFEPNLPYYID 60
 QY 61 GNVKLTQSAIIRYADKNNMLGCGPKERAETSMLEGAVALDIRYGVSRATYSKDFETLV 120
 DB 61 GNVKLTQSAIIRYADKNNMLGCGPKERAETSMLEGAVALDIRYGVSRATYSKDFETLV 120
 QY 121 DFLSKPEMLKMEFRLCHKTYLNGDHYHNPMLYDAVDVLYMPPKLDLAFPKLVCKR 180
 DB 121 DFLSKPEMLKMEFRLCHKTYLNGDHYHNPMLYDAVDVLYMPPKLDLAFPKLVCKR 180
 QY 181 KRLEIAPIQDKLKSSKYTAMPLOGWQATFGGADHPPKSDLYPRGSRNVRDSSGRAEIT 240
 DB 181 KRLEIAPIQDKLKSSKYTAMPLOGWQATFGGADHPPKSDLYPRGSRNVRDSSGRAEIT 240
 QY 241 RILPVKGSKLRALEHGLEDFLOKQOYGISKYSGFVAYTD-CTESGQNLCLCGSNV 299
 DB 241 RILPVKGSKLRALEHGLEDFLOKQOYGISKYSGFVAYTD-CTESGQNLCLCGSNV 299
 QY 229 -IP---gsitraassdhal-----gtlrsdnakepdydc----- 260
 DB 229 -IP---gsitraassdhal-----gtlrsdnakepdydc----- 260
 QY 300 CGGKNCILGSDGK 314
 DB 300 CGGKNCILGSDGK 314
 QY 261 cgnqn-----scgrk 270
 DB 261 cgnqn-----scgrk 270

AC Y52288;
 XX 09-FEB-2000 (first entry)
 DT Human HCP/GST fusion protein.
 DE
 XX
 XX
 KW Haematopoietic cell phosphatase; HCP; cloning; expression;
 KW epithelial cell; growth factor receptor; interleukin-3; IL-3;
 KW erythropoietin; EPO; negative regulator; signal transduction;
 KW modulation; activity; inhibitor; stimulation; cytokine therapy;
 KW antisense therapy; gene therapy; treatment; diagnosis;
 KW glutathione-S-transferase; fusion.
 XX
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 XX
 PN WO954450-A1.
 PD 28-OCT-1999.
 XX
 XX
 PF 15-APR-1999; 99WO-US08228.
 XX
 XX
 PR 17-APR-1998; 98US-0082072.
 XX
 XX
 PA (ORTH) ORTHO-MCNEIL PHARM INC.
 XX
 XX
 PI Jolliffe L, Barbone F, Shanahan M, Xu D;
 DR WPI: 2000-013244/01.
 DR N-PSDB: Z32916.
 XX
 PT Assay methods for modulators of hematopoietic cell phosphatase -
 PT Claim 3; Fig 4B; 57pp; English.
 PS
 XX
 XX
 CC This sequence represents a full-length human haematopoietic
 CC cell phosphatase (HCP)/glutathione-S-transferase (GST) fusion
 CC protein. This was expressed in E. coli and subsequently
 CC purified for use in a variety of activity assays.
 CC HCP is a 68 kD non-membrane associated protein found primarily in
 CC haematopoietic cells, and to some extent in epithelial cells. It
 CC has been shown to associate with several growth factor receptors
 CC following ligand stimulation and acts as a negative regulator
 CC of receptor activation, and hence of haematopoietic signal transduction.
 CC The receptors it regulates include the interleukin-3 (IL-3) and
 CC erythropoietin (EPO) receptors. The invention relates to a novel
 CC method of identifying compounds that modulate HCP activity. This
 CC comprises combining a modulator of HCP protein activity with HCP protein
 CC and HCP protein substrate, and measuring an effect of the modulator on
 CC the activity of the HCP protein. Inhibitors of HCP can be used to
 CC modulate the activity of haematopoietic receptors. HCP has been shown to
 CC be a negative regulator of one or more signal transduction pathways in
 CC haematopoietic cells, and the identification of a HCP inhibitor would
 CC provide a synthetic stimulator to the haematopoietic system could be
 CC used in conjunction with other cytokine therapy, e.g., administration of
 CC EPO. HCP is also associated with several growth factors found in
 CC haematopoietic cells and it is possible that a HCP inhibitor may also
 CC function as a haematopoietic cell potentiator, enhancing the effect of
 CC growth factors, decreasing the necessary dose of the factor. The HCP
 CC polynucleotides are sources of probes and primers, and can be used to
 CC design antisense sequences, and in gene therapy. The protein, or
 CC active fragments thereof, may be used to treat or diagnose disorders
 CC which require the modulation of the HCP activity. Small molecules that
 CC inhibit the action of HCP can also augment the effect of EPO, which
 CC meaning that a lower therapeutic dose of EPO may be utilised.
 CC
 XX
 XX
 SQ Sequence 824 AA:
 Query Match 66.7%; Score 1233.5; DB 21; Length 824;
 Best Local Similarity 81.9%; Pred. No. 1.3e-100;
 Matches 245; Conservative 7; Mismatches 24; Indels 23; Gaps 6;

PT antibodies, used to deliver therapeutic or diagnostic agents to, or
 PT through, the gastrointestinal tract, e.g. insulin or leuprolide
 PS Example: Page 226; 294pp; English.

CC This sequence represents a fusion protein between
 CC glutathione-S-transferase and a gastro-intestinal transport protein
 CC binding peptide. The invention relates to purified proteins (I) that bind
 CC specifically to at least one of the gastro-intestinal (GI) tract
 CC receptors human intestinal peptide-associated transporter (HPT1),
 CC hPEPT1, D2H and human sucrose-isomaltase complex (HSI). (I) provide
 CC active transport of therapeutic agents through human and animal GI tissue
 CC (into the blood) for in vivo delivery, particularly for treatment or
 CC prevention of hypertension, diabetes, osteoporosis, haemophilia, anaemia,
 CC cancer, migraine, or angina pectoris. Specifically they are used to
 CC deliver insulin or leuprolide, but many other suitable therapeutic agents
 CC are disclosed, including genes or inhibitory nucleic acid, imaging agents
 CC and antigens. (I) may also provide targeting to the GI tract. Other uses
 CC of (I) are: (i) to determine the level of specified receptors in a sample
 CC (in a binding assay); and (ii) to screen for molecules that bind (I).
 CC Immunogenic analogues or derivatives of (I) are used to raise antibodies
 CC and in immunoassays. The antibodies are used to locate, detect and
 CC measure (I), e.g. for imaging, monitoring treatment, tissue analysis
 CC etc., also for peptide purification and immobilisation.

XX Sequence 282 AA:

Query Match 66.6%; Score 1232.5; DB 20; Length 282;
 Best Local Similarity 94.7%; Pred. No. 4.1e-101;
 Matches 233; Conservative 2; Mismatches 6; Indels 5; Gaps 2;

OY 1 MSPILGYWKIKGLVOPTRLLLEYLEEKYEHEHLYERDEGDKWRKKFELGLEPNNLPYYID 60
 DB 1 MSPILGYWKIKGLVOPTRLLLEYLEEKYEHEHLYERDEGDKWRKKFELGLEPNNLPYYID 60
 OY 61 GGVKLTQSMATIRYIADKHNMLGCGCKERAETISMEGAVLDIRYGSRIAYSKDFETLVY 120
 DB 61 GGVKLTQSMATIRYIADKHNMLGCGCKERAETISMEGAVLDIRYGSRIAYSKDFETLVY 120
 OY 121 DFLSKLPEMLKMPEDRLCHRTYTLNGDHVTHPDMLVDALDVLVYMPMCDLAPKLVCFK 180
 DB 121 DFLSKLPEMLKMPEDRLCHRTYTLNGDHVTHPDMLVDALDVLVYMPMCDLAPKLVCFK 180
 OY 181 KRLEAIPOIDKYLKSSKYIAMPLOGQATFGGDHPKPSDLVPRGSPNSRVSNGRAEIT 240
 DB 181 KRLEAIPOIDKYLKSSKYIAMPLOGQATFGGDHPKPSDLVPRGSPNSRVSNGRAEIT 240
 OY 241 RIPLVYK 246
 DB 239 r---yk 241

RESULT 13
 W74113 standard; peptide: 248 AA.

AC W74113:
 XX 04-MAY-1999 (first entry)

DE GST/GI transport receptor binding protein fusion sequence.
 XX Gastro-intestinal transport receptor; binding protein; hsi; hPT1;
 XX D2H; hPEPT1; human; GI tract receptor; sucrose-isomaltase complex;
 KW intestinal peptide-associated transporter; hypertension; diabetes;
 KW osteoporosis; haemophilia; anaemia; cancer; migraine; angina pectoris;
 KW therapeutic agent delivery; therapy; fusion protein.

OS Homo sapiens.
 XX
 PN W09051325-A2.
 XX

PD 19-NOV-1998.
 XX
 PF 15-MAY-1998: 98WO-US10088.
 XX
 PR 15-MAY-1997: 97US-0046595.

PA (CYTO-) CYTOGEN CORP.
 PA (ELAN-) ELAN CORP PLC.
 PI Alvarez VL, Belinka BA, Cagney GW, Carter JM, Lambkin IJ;
 PI Omahony DJ, Patterson CA, Singleton J;
 DR WPI: 1999-009568/01.

PT New proteins that bind specifically to receptors in the
 PT gastro-intestinal tract and related nucleic acid - chimaeras and
 PT antibodies, used to deliver therapeutic or diagnostic agents to, or
 PT through, the gastrointestinal tract, e.g. insulin or leuprolide

PS Example: Page 216; 294pp; English.

CC This sequence represents a fusion protein between
 CC glutathione-S-transferase and a gastro-intestinal transport protein
 CC binding peptide. The invention relates to purified proteins (I) that bind
 CC specifically to at least one of the gastro-intestinal (GI) tract
 CC receptors human intestinal peptide-associated transporter (HPT1),
 CC hPEPT1, D2H and human sucrose-isomaltase complex (HSI). (I) provide
 CC active transport of therapeutic agents through human and animal GI tissue
 CC (into the blood) for in vivo delivery, particularly for treatment or
 CC prevention of hypertension, diabetes, osteoporosis, haemophilia, anaemia,
 CC cancer, migraine, or angina pectoris. Specifically they are used to
 CC deliver insulin or leuprolide, but many other suitable therapeutic agents
 CC are disclosed, including genes or inhibitory nucleic acid, imaging agents
 CC and antigens. (I) may also provide targeting to the GI tract. Other uses
 CC of (I) are: (i) to determine the level of specified receptors in a sample
 CC (in a binding assay); and (ii) to screen for molecules that bind (I).
 CC Immunogenic analogues or derivatives of (I) are used to raise antibodies
 CC and in immunoassays. The antibodies are used to locate, detect and
 CC measure (I), e.g. for imaging, monitoring treatment, tissue analysis
 CC etc., also for peptide purification and immobilisation.

XX Sequence 248 AA:

Query Match 66.6%; Score 1232; DB 20; Length 248;
 Best Local Similarity 95.9%; Pred. No. 3.9e-101;
 Matches 231; Conservative 2; Mismatches 6; Indels 2; Gaps 1;

OY 1 MSPILGYWKIKGLVOPTRLLLEYLEEKYEHEHLYERDEGDKWRKKFELGLEPNNLPYYID 60
 DB 1 MSPILGYWKIKGLVOPTRLLLEYLEEKYEHEHLYERDEGDKWRKKFELGLEPNNLPYYID 60
 OY 61 GGVKLTQSMATIRYIADKHNMLGCGCKERAETISMEGAVLDIRYGSRIAYSKDFETLVY 120
 DB 61 GGVKLTQSMATIRYIADKHNMLGCGCKERAETISMEGAVLDIRYGSRIAYSKDFETLVY 120
 OY 121 DFLSKLPEMLKMPEDRLCHRTYTLNGDHVTHPDMLVDALDVLVYMPMCDLAPKLVCFK 180
 DB 121 DFLSKLPEMLKMPEDRLCHRTYTLNGDHVTHPDMLVDALDVLVYMPMCDLAPKLVCFK 180
 OY 181 KRLEAIPOIDKYLKSSKYIAMPLOGQATFGGDHPKPSDLVPRGSPNSRVSNGRAEIT 240
 DB 181 KRLEAIPOIDKYLKSSKYIAMPLOGQATFGGDHPKPSDLVPRGSPNSRVSNGRAEIT 240
 OY 241 R 241
 DB 239 r 239

RESULT 14
 W74122 standard; peptide: 248 AA.
 ID W74122
 XX

AC W74122;
 XX
 DT 04-MAY-1999 (first entry)
 XX
 DE GST/GI transport receptor binding protein fusion sequence.
 XX
 KM Gastro-intestinal transport receptor; binding protein; hSI; hPRL;
 KM D2H; hPEPTL; human; GI tract receptor; sucrose-isomaltase complex;
 KM intestinal peptide-associated transporter; hypertension; diabetes;
 KM osteoporosis; haemophilia; anaemia; cancer; migraine; angina pectoris;
 KM therapeutic agent delivery; therapy; fusion protein.
 XX
 OS Homo sapiens.
 XX
 PN WO9851325-A2.
 XX
 PD 19-NOV-1998.
 XX
 PF 15-MAY-1998; 98WO-US10088.
 XX
 PR 15-MAY-1997; 97US-0046595.
 XX
 PA (CYTO-) CYTOGEN CORP.
 PA (ELAN-) ELAN CORP PLC.
 XX
 PI Alvarez VL, Belinka BA, Cagney GM, Carter JM, Lambkin IJ;
 PI Omahony DJ, Patterson CA, Singleton J;
 XX
 DR WPI: 1999-009568/01.
 XX
 PT New proteins that bind specifically to receptors in the
 PT gastro-intestinal tract and related nucleic acid - chimeras and
 PT antibodies, used to deliver therapeutic or diagnostic agents to, or
 PT through, the gastrointestinal tract, e.g. insulin or leuprolide
 XX
 PS Example: Page 222-223; 294pp; English.
 XX
 CC This sequence represents a fusion protein between
 CC glutathione-S-transferase and a gastro-intestinal transport protein
 CC binding peptide. The invention relates to purified proteins (I) that bind
 CC specifically to at least one of the gastro-intestinal (GI) tract
 CC receptors human intestinal peptide-associated transporter (hPRL),
 CC hPEPTL, D2H and human sucrose-isomaltase complex (hSI). (I) provide
 CC active transport of therapeutic agents through human and animal GI tissue
 CC (into the blood) for in vivo delivery, particularly for treatment or
 CC prevention of hypertension, diabetes, osteoporosis, haemophilia, anaemia,
 CC cancer, migraine, or angina pectoris. Specifically they are used to
 CC deliver insulin or leuprolide, but many other suitable therapeutic agents
 CC are disclosed, including genes or inhibitory nucleic acid, imaging agents
 CC and antigens. (I) may also provide targeting to the GI tract. Other uses
 CC of (I) are: (i) to determine the level of specified receptors in a sample
 CC (in a binding assay); and (ii) to screen for molecules that bind (I).
 CC Immunogenic analogues or derivatives of (I) are used to raise antibodies
 CC and in immunoassays. The antibodies are used to locate, detect and
 CC measure (I), e.g. for imaging, monitoring treatment, tissue analysis
 CC etc., also for peptide purification and immobilisation.
 CC
 XX Sequence 248 AA;
 SQ

Query Match 66.6%; Score 1232; DB 20; Length 248;
 Best Local Similarity 95.9%; Pred. No. 3,9e-10;
 Matches 231; Conservative 2; Mismatches 6; Indels 2; Gaps 1;

QY 1 MSPIIGYWKIKLVQPTRLLEYLEEKYEHEHYERDEGDKMKKFEELGEPNLPLYTD 60
 DB 1 MSPILGYWKIKLVQPTRLLEYLEEKYEHEHYERDEGDKMKKFEELGEPNLPLYTD 60
 QY 61 GGVKLTQSMATIRYADKNNMLGCPKERAETSMLEGAVALDRYGVSRAYSKDEFTTKV 120
 DB 61 ggvkltqsmatiryladknnmlgcpkeraetasmlegavaldrrygvsrayskdefttkv 120
 QY 121 DFLSKLPKMLKMFEDRLCHRTYLNQGDVTHPDPMLYDALDVVLYMDPKCLAFPKLVCFK 180

DB 121 dflsklpemlkmfedrlchrtcylnqgdvthpdpmlfydalvlymdpncldafklycfk 180
 QY 181 KRIFAIPQIDKYLKSSKSIAMPLOGMOTFEGGCHPPEKSDLPVPGSPRSRDSSGRAEIT. 240
 DB 181 krifaipqidkylkssksiyawplogmotfeggchppesdldvprgspg-1pgsttraas 238
 QY 241 R 241
 DB 239 r 239

RESULT 15
 W74117
 ID W74117 standard; peptide: 258 AA.
 XX
 AC W74117;
 XX
 DT 04-MAY-1999 (first entry)
 XX
 DE GST/GI transport receptor binding protein fusion sequence.
 XX
 KM Gastro-intestinal transport receptor; binding protein; hSI; hPRL;
 KM D2H; hPEPTL; human; GI tract receptor; sucrose-isomaltase complex;
 KM intestinal peptide-associated transporter; hypertension; diabetes;
 KM osteoporosis; haemophilia; anaemia; cancer; migraine; angina pectoris;
 KM therapeutic agent delivery; therapy; fusion protein.
 XX
 OS Homo sapiens.
 XX
 PN WO9851325-A2.
 XX
 PD 19-NOV-1998.
 XX
 PF 15-MAY-1998; 98WO-US10088.
 XX
 PR 15-MAY-1997; 97US-0046595.
 XX
 PA (CYTO-) CYTOGEN CORP.
 PA (ELAN-) ELAN CORP PLC.
 XX
 PI Alvarez VL, Belinka BA, Cagney GM, Carter JM, Lambkin IJ;
 PI Omahony DJ, Patterson CA, Singleton J;
 XX
 DR WPI: 1999-009568/01.
 XX
 PT New proteins that bind specifically to receptors in the
 PT gastro-intestinal tract and related nucleic acid - chimeras and
 PT antibodies, used to deliver therapeutic or diagnostic agents to, or
 PT through, the gastrointestinal tract, e.g. insulin or leuprolide
 XX
 PS Example: Page 219; 294pp; English.
 XX
 CC This sequence represents a fusion protein between
 CC glutathione-S-transferase and a gastro-intestinal transport protein
 CC binding peptide. The invention relates to purified proteins (I) that bind
 CC specifically to at least one of the gastro-intestinal (GI) tract
 CC receptors human intestinal peptide-associated transporter (hPRL),
 CC hPEPTL, D2H and human sucrose-isomaltase complex (hSI). (I) provide
 CC active transport of therapeutic agents through human and animal GI tissue
 CC (into the blood) for in vivo delivery, particularly for treatment or
 CC prevention of hypertension, diabetes, osteoporosis, haemophilia, anaemia,
 CC cancer, migraine, or angina pectoris. Specifically they are used to
 CC deliver insulin or leuprolide, but many other suitable therapeutic agents
 CC are disclosed, including genes or inhibitory nucleic acid, imaging agents
 CC and antigens. (I) may also provide targeting to the GI tract. Other uses
 CC of (I) are: (i) to determine the level of specified receptors in a sample
 CC (in a binding assay); and (ii) to screen for molecules that bind (I).
 CC Immunogenic analogues or derivatives of (I) are used to raise antibodies
 CC and in immunoassays. The antibodies are used to locate, detect and
 CC measure (I), e.g. for imaging, monitoring treatment, tissue analysis
 CC etc., also for peptide purification and immobilisation.
 CC
 XX

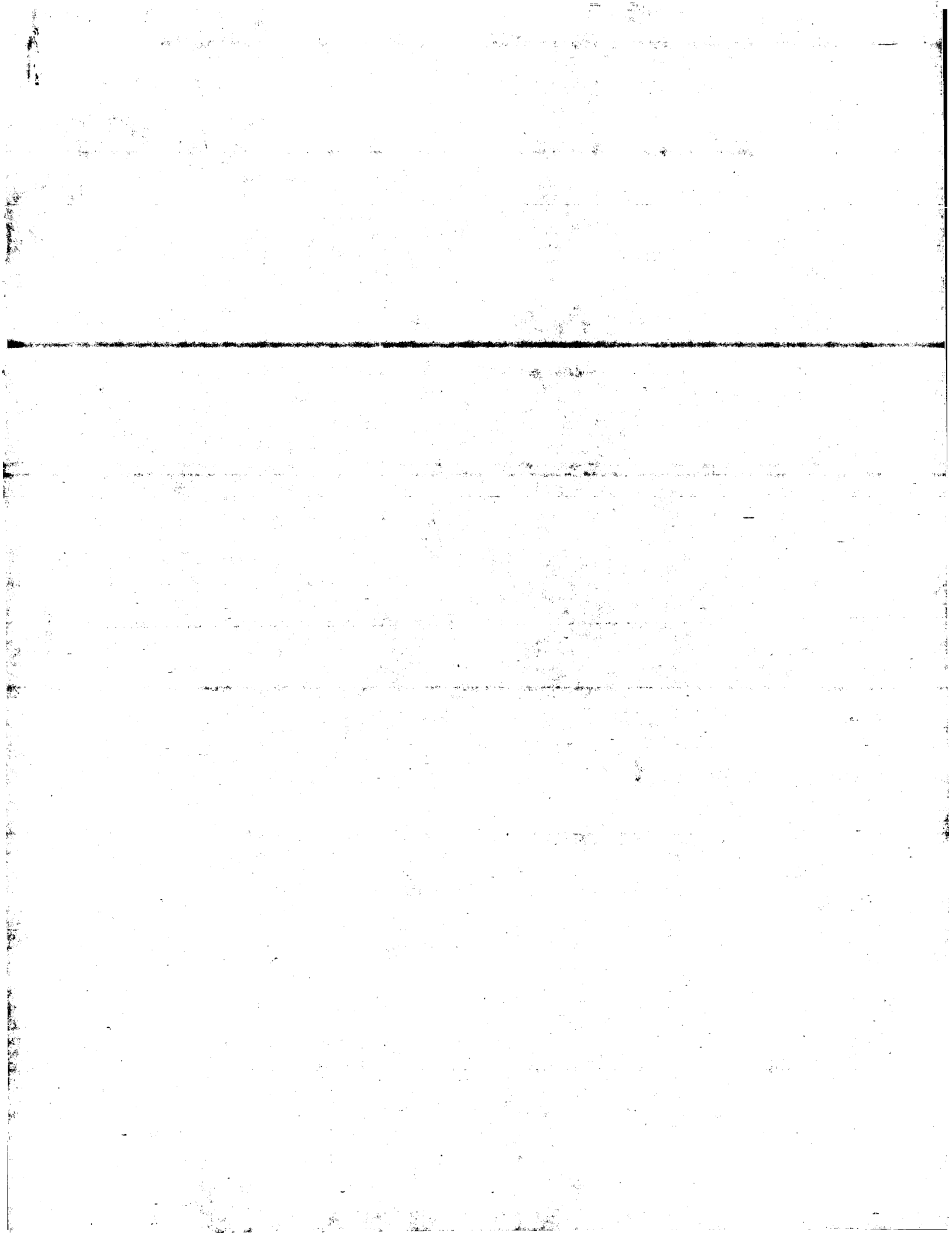
SQ Sequence 258 AA:

Query Match 66.6%; Score 1232; DB 20; Length 258;

Best Local Similarity 95.1%; Pred. No. 4, 1e-101; Matches 231; Conservative 2; Mismatches 8; Indels 2; Gaps 1;

QY 1 MSPILGYWKIKGLVOPTRLLLEYLAEKYEENLYERDEGDKWRNKKFELGLEPPNLPYYID 60
 |||
 Db 1 mspilgywkikglvqptrllleyleekyeenhlyerdegdkwrnkkelglfepnlpyyid 60
 QY 61 GPKVLTQSMATIRYIADKNNMLGCGPKERAHISMLEGAVLDIRYGVSRITAYSKDEFTLKV 120
 |||
 Db 61 gpkvltqsmatiryladknnmlgscpkeraeismllegavldirygvsritayskdfetlkv 120
 QY 121 DFLSKLPEMLKMFEDRLCHKTYLNGDHTHPPEMLYDALDVLYMDPMCLDAFPKLVCFK 180
 |||
 Db 121 dflsklpemlkmfedrlchkyngdhvchpdtlmdalvvlvymdpmcldatpklvcfk 180
 QY 181 KRLEAIPOIDKYLLKSSKYIAMPLOGWQATFGGDDHPKSDLVPRGSPNSRVDSSGRAEIT 240
 |||
 Db 181 krleaipqidkylksskylawplgwgatfaggdhpkksdlvprgspnsvdssgraaas 238
 QY 241 RIP 243
 Db 239 stp 241

Search completed: March 18, 2001, 05:32:01
 Job time: 2079 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2001, 04:33:51 ; Search time 65.51 Seconds
(without alignments)
94.020 Million cell updates/sec

Title: US-09-402-488a-2

Perfect score: 1850
Sequence: 1 MSPILGYWKIKGLVQPTRLLE.....PKQSHNDGDFEELPEEYLQ 343

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/PCRTUS.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1234.5	66.7	354	2	US-08-216-894-4
2	1234.5	66.7	538	3	US-08-974-549A-602
3	1234.5	66.7	564	2	US-08-216-894-2
4	1234.5	66.7	643	2	US-08-216-894-8
5	1232.5	66.6	515	3	US-08-974-549A-604
6	1229.5	66.5	517	3	US-08-974-549A-606
7	1229	66.4	530	3	US-08-974-549A-603
8	1227	66.3	232	4	PCR-US94-09700-6
9	1226	66.3	397	4	PCR-US94-09700-11
10	1226	66.3	514	3	US-08-974-549A-605
11	1224	66.2	579	2	US-08-864-224-11
12	1222	66.1	352	1	US-08-395-507-1
13	1219.5	65.9	426	3	US-08-737-248-4
14	1219	65.9	362	1	US-08-395-507-2
15	1207	65.2	291	1	US-08-102-757-9
16	1206.5	65.2	472	2	US-08-216-894-10
17	1206	65.2	435	4	PCR-US95-04439-1
18	1205	65.1	647	2	US-08-305-764C-56
19	1201.5	64.9	536	3	US-08-974-180-20
20	1189	64.3	291	1	US-08-102-757-11
21	1189	64.3	307	2	US-08-216-894-6
22	1187	64.2	616	3	US-08-895-707-2
23	1179	63.7	218	3	US-08-470-837-26
24	471.5	25.5	217	3	US-08-297-431B-33
25	378	20.4	134	1	US-07-854-596B-9
26	375	20.3	483	1	US-07-854-596B-47
27	369	19.9	82	1	US-08-715-252-2
28	369	19.9	82	2	US-08-453-051-4

29	368	19.9	92	1	US-08-186-222-4	Sequence 4, Appl
30	367	19.8	65	1	US-07-970-596-1	Sequence 1, Appl
31	367	19.8	65	1	US-07-763-860-1	Sequence 1, Appl
32	367	19.8	65	1	US-07-854-596B-2	Sequence 2, Appl
33	367	19.8	65	1	US-08-058-699-1	Sequence 1, Appl
34	367	19.8	65	1	US-07-910-528-3	Sequence 3, Appl
35	367	19.8	65	1	US-08-348-972-3	Sequence 3, Appl
36	367	19.8	65	1	US-08-255-272-17	Sequence 17, Appl
37	367	19.8	65	1	US-08-406-948A-6	Sequence 6, Appl
38	367	19.8	65	1	US-08-367-758B-14	Sequence 14, Appl
39	367	19.8	65	2	US-08-909-735-14	Sequence 14, Appl
40	367	19.8	65	5	5164304-10	Patent No. 5164304
41	367	19.8	65	5	5167960-1	Patent No. 5167960
42	367	19.8	66	5	5422249-2	Patent No. 5422249
43	367	19.8	483	1	US-07-854-596B-43	Sequence 43, Appl
44	362	19.6	64	1	US-07-763-860-2	Sequence 2, Appl
45	360	19.5	65	1	US-08-199-506A-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-216-894-4
Sequence 4, Application US/08216894
Patent No. 5876734
GENERAL INFORMATION:
APPLICANT: Kirchhoff, Louis V.
TITLE OF INVENTION: POLYPEPTIDES FOR DIAGNOSING INFECTION
TITLE OF INVENTION: WITH TRYPAOSOMA CRUZI
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/216,894
FILING DATE: 24-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 85326/102/DRLO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 354 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-216-894-4

Query Match 66.7%; Score 1234.5; DB 2; Length 354;
Best Local Similarity 89.3%; Pred. No. 9.2e-106;
Matches 233; Conservative 10; Mismatches 9; Indels 9; Gaps 2;

QY 1 MSPILGYWKIKGLVQPTRLLEYLEEKYEHELYERDEGDKWRKKFELGLEPNNLPYYID 60
|||||
DB 1 MSPILGYWKIKGLVQPTRLLEYLEEKYEHELYERDEGDKWRKKFELGLEPNNLPYYID 60
|||||
QY 61 GDVKTQSMATIRYIADKHNMLGGCKERAETISMLGCAVLDIRYGVSRAYSQDFELRV 120
|||||

Db 61 GDVKLTQSMALIRYADKHNMLGCGPKERAETISMEGAVLDIRGVSRIVASKDEFTLKV 120
QY 121 DFLSLPEMLKMFEDRLCHKTYLNDHVTHPDPMLYDALDVLYNDPCLDAFPLVCEK 180
Db 121 DFLSLPEMLKMFEDRLCHKTYLNDHVTHPDPMLYDALDVLYNDPCLDAFPLVCEK 180
QY 181 KRIEAIPOIDKYLKSSKTYIAMPLOGWATFGGSDHPKSDLVPRGSPNSRYDSSGRAEIT 240
Db 181 KRIEAIPOIDKYLKSSKTYIAMPLOGWATFGGSDHPKSDLVPRGSPNSRYDSSGRAEIT 240
QY 236 RAEITRIPLYKGSILKALKE 256
Db 241 KKEMTKL---REKYKAEKE 257

RESULT 2
US-08-974-549A-602
Sequence 602, Application US/08974549A
Patent No. 6166178
GENERAL INFORMATION:
APPLICANT: Cecch, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
CLASSIFICATION: 536
PRIOR APPLICATION NUMBER:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-00261005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 602:
SEQUENCE CHARACTERISTICS:
LENGTH: 538 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..538
OTHER INFORMATION: glutathione-S-transferase sequence,
OTHER INFORMATION: thrombin cleavage sequence, recognition
OTHER INFORMATION: sequence for heart muscle protein
OTHER INFORMATION: kinase, residues introduced by cloning
OTHER INFORMATION: and hTERT protein fragment"
US-08-974-549A-602

Query Match 66.7%; Score 1234.5; DB 3; Length 538;
Best Local Similarity 85.1%; Pred. No. 1.6e-105;
Matches 240; Conservative 6; Mismatches 13; Indels 23; Gaps 4;

QY 1 MSPILGYWKIKIGVOPITLLEYLEEKEEHLYEDEDECKNNKFEELGEPNLPYID 60
Db 1 MSPILGYWKIKIGVOPITLLEYLEEKEEHLYEDEDECKNNKFEELGEPNLPYID 60
QY 61 GDVKLTQSMALIRYADKHNMLGCGPKERAETISMEGAVLDIRGVSRIVASKDEFTLKV 120
Db 61 GDVKLTQSMALIRYADKHNMLGCGPKERAETISMEGAVLDIRGVSRIVASKDEFTLKV 120
QY 121 DFLSLPEMLKMFEDRLCHKTYLNDHVTHPDPMLYDALDVLYNDPCLDAFPLVCEK 180
Db 121 DFLSLPEMLKMFEDRLCHKTYLNDHVTHPDPMLYDALDVLYNDPCLDAFPLVCEK 180
QY 181 KRIEAIPOIDKYLKSSKTYIAMPLOGWATFGGSDHPKSDLVPRGSPNSRYDSSGRAEIT 240
Db 181 KRIEAIPOIDKYLKSSKTYIAMPLOGWATFGGSDHPKSDLVPRGSPNSRYDSSGRAEIT 240
QY 241 RIP-----LYKG---KSLRKALKEHGL-----EDFL 264
Db 236 KIPGSIITSLTLCISLCYGDMEKIFAGIRDGLLRVDDFL 277

RESULT 3
US-08-216-894-2
Sequence 2, Application US/08216894
Patent No. 5876734
GENERAL INFORMATION:
APPLICANT: Kirchhoff, Louis V.
APPLICANT: Otsu, Keiko
TITLE OF INVENTION: POLYPEPTIDES FOR DIAGNOSING INFECTION
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/216,894
FILING DATE: 24-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 85326/102/DRLO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 564 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-216-894-2

Query Match 66.7%; Score 1234.5; DB 2; Length 564;
Best Local Similarity 89.3%; Pred. No. 1.7e-105;
Matches 233; Conservative 10; Mismatches 9; Indels 9; Gaps 2;

QY 1 MSPILGYWKIKGLVOPTRILLEYLEEKYEDEHLYERDEGDKWNRKKEFGLEFPNLPYYID 60
DB 1 MSPILGYWKIKGLVOPTRILLEYLEEKYEDEHLYERDEGDKWNRKKEFGLEFPNLPYYID 60
QY 61 GGVKLTQSMALIRYIADKNNMGGCPKERAETSMLEGAVALDIRYGSRTAYSDFETLKY 120
DB 61 GGVKLTQSMALIRYIADKNNMGGCPKERAETSMLEGAVALDIRYGSRTAYSDFETLKY 120
QY 121 DFLSKLPEMLKMFEDRLCKHTYLNAGDHVTHPPDMLDALDVLYMPKCLDAFPKLYCKR 180
DB 121 DFLSKLPEMLKMFEDRLCKHTYLNAGDHVTHPPDMLDALDVLYMPKCLDAFPKLYCKR 180
QY 181 KRIEALPQIDKYLKSSKRYIAMPLOGQATFGGDHPKSDLYVRGSPNSRVD-----SSG 235
DB 181 KRIEALPQIDKYLKSSKRYIAMPLOGQATFGGDHPKSDLYVRGSPNSRVD-----SSG 235
QY 236 RAEITRIPLYKSKSLKALKE 256
DB 241 KKEMTKL---REKVKAEKE 257

RESULT 4
US-08-216-894-8
Sequence 8, Application US/08216894
Patent No. 5876734
GENERAL INFORMATION:
APPLICANT: Kirchoff, Louis V.
APPLICANT: Otsu, Keiko
TITLE OF INVENTION: POLYPEPTIDES FOR DIAGNOSING INFECTION
TITLE OF INVENTION: WITH TRYPAOSOMA CRUZI
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/216,894
FILING DATE: 24-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 85326/102/DRLO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 643 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-216-894-8

Query Match 66.7%; Score 1234.5; DB 2; Length 643;
Best Local Similarity 89.3%; Pred. No. 2e-105;
Matches 233; Conservative 10; Mismatches 9; Indels 9; Gaps 2;

QY 1 MSPILGYWKIKGLVOPTRILLEYLEEKYEDEHLYERDEGDKWNRKKEFGLEFPNLPYYID 60
DB 1 MSPILGYWKIKGLVOPTRILLEYLEEKYEDEHLYERDEGDKWNRKKEFGLEFPNLPYYID 60
QY 61 GGVKLTQSMALIRYIADKNNMGGCPKERAETSMLEGAVALDIRYGSRTAYSDFETLKY 120
DB 61 GGVKLTQSMALIRYIADKNNMGGCPKERAETSMLEGAVALDIRYGSRTAYSDFETLKY 120
QY 121 DFLSKLPEMLKMFEDRLCKHTYLNAGDHVTHPPDMLDALDVLYMPKCLDAFPKLYCKR 180
DB 121 DFLSKLPEMLKMFEDRLCKHTYLNAGDHVTHPPDMLDALDVLYMPKCLDAFPKLYCKR 180
QY 181 KRIEALPQIDKYLKSSKRYIAMPLOGQATFGGDHPKSDLYVRGSPNSRVD-----SSG 235
DB 181 KRIEALPQIDKYLKSSKRYIAMPLOGQATFGGDHPKSDLYVRGSPNSRVD-----SSG 235
QY 236 RAEITRIPLYKSKSLKALKE 256
DB 241 KKEMTKL---REKVKAEKE 257

RESULT 5
US-08-974-549A-604
Sequence 604, Application US/08974549A
Patent No. 6166178
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Langer, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 604:
SEQUENCE CHARACTERISTICS:
LENGTH: 515 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..515
OTHER INFORMATION: /note= "Fusion protein composed of
OTHER INFORMATION: glutathione-S-transferase sequence,
OTHER INFORMATION: thrombin cleavage sequence, recognition
OTHER INFORMATION: sequence for heart muscle protein
OTHER INFORMATION: kinase, residues introduced by cloning
OTHER INFORMATION: and hprt protein fragment"
US-08-974-549A-604

Query Match 66.6%; Score 1232.5; DB 3; Length 515;
Best Local Similarity 92.4%; Pred. No. 2.3e-105;
Matches 232; Conservative 6; Mismatches 8; Indels 5; Gaps 1;
QY 1 MSPILGWKIKGLVOPTRILLEYLEERKYEHELYERDEGCKWRNKKFELGLEPNIPLYID 60
DB 1 MSPILGWKIKGLVOPTRILLEYLEERKYEHELYERDEGCKWRNKKFELGLEPNIPLYID 60
QY 61 GDVKLTQSMATIRYIADKNHMLGGCKREKREIEMLEGAVLIDITYGSRIRAYSDPEFLKY 120
DB 61 GDVKLTQSMATIRYIADKNHMLGGCKREKREIEMLEGAVLIDITYGSRIRAYSDPEFLKY 120
QY 121 DFLSKLPEMKMFEDRCHKTYINGDVHPDPMYDALDVLYLMDPMCLDAPPKLVCFK 180
DB 121 DFLSKLPEMKMFEDRCHKTYINGDVHPDPMYDALDVLYLMDPMCLDAPPKLVCFK 180
QY 181 KRTEAIPQIDKYLKSSKIYAMPLOGWQATEGGGDHPKSDLVPRGSPNSRVSSGRAEIT 240
DB 181 KRTEAIPQIDKYLKSSKIYAMPLOGWQATEGGGDHPKSDLVPRGSPNSRVSSGRAEIT 240
QY 241 RIPLYKGSJR 251

Db 236 KMSYVVELLR 246
RESULT 6
US-08-974-549A-606
Sequence 606, Application US/08974549A
Patent No. 6166178
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: Human telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 606:
SEQUENCE CHARACTERISTICS:

LENGTH: 517 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..517
OTHER INFORMATION: /note="fusion protein composed of
OTHER INFORMATION: glutathione-S-transferase sequence,
OTHER INFORMATION: thrombin cleavage sequence, recognition
OTHER INFORMATION: sequence for heart muscle protein
OTHER INFORMATION: kinase residues introduced by cloning
and hprt protein fragment"
US-08-974-549A-606

Query Match 66.5%; Score 1229.5; DB 3; Length 517;
Best Local Similarity 94.7%; Pred. No. 4.4e-105;
Matches 230; Conservative 4; Mismatches 4; Indels 5; Gaps 1;

QY 1 MSPILGYWKIKGLVQPTRLLEYLEEKYEHEHLYERDEGDKWRKKFELGLEFPNLPYYID 60
DB 1 MSPILGYWKIKGLVQPTRLLEYLEEKYEHEHLYERDEGDKWRKKFELGLEFPNLPYYID 60
QY 61 GDVKLTQSMALIRYIADKHNMLGCGCPKERAETISMLEGAVLDIRYGVSRAYSXDFETLVK 120
DB 61 GDVKLTQSMALIRYIADKHNMLGCGCPKERAETISMLEGAVLDIRYGVSRAYSXDFETLVK 120
QY 121 DFLSKLPEMLKMFEDRLCKHTYINGDHVTHPPDMLYDALDVLVYMPMCLDAPFKLVCKR 180
DB 121 DFLSKLPEMLKMFEDRLCKHTYINGDHVTHPPDMLYDALDVLVYMPMCLDAPFKLVCKR 180
QY 181 KRLEAIPOIDKYLKSSKTYIAMPLOGQATFGGSDHPKSDLVPRGSPNSRVDS--SRAETI 240
DB 181 KRLEAIPOIDKYLKSSKTYIAMPLOGQATFGGSDHPKSDLVPRGSPNSRVDS--SRAETI 240
QY 241 RIP 243
DB 236 KMP 238

RESULT 7
US-08-974-549A-603
Sequence 603, Application US/08974549A
Patent No. 6166178
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-00261005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 603:
SEQUENCE CHARACTERISTICS:
LENGTH: 530 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..530
OTHER INFORMATION: /note="fusion protein composed of
OTHER INFORMATION: glutathione-S-transferase sequence,
OTHER INFORMATION: thrombin cleavage sequence, recognition
OTHER INFORMATION: sequence for heart muscle protein
OTHER INFORMATION: kinase residues introduced by cloning,
OTHER INFORMATION: eight consecutive His residues and hprt
protein fragment"
US-08-974-549A-603

Query Match 66.4%; Score 1229; DB 3; Length 530;
Best Local Similarity 90.3%; Pred. No. 5e-105;
Matches 232; Conservative 6; Mismatches 13; Indels 6; Gaps 1;

QY 1 MSPILGYWKIKGLVQPTRLLEYLEEKYEHEHLYERDEGDKWRKKFELGLEFPNLPYYID 60
DB 1 MSPILGYWKIKGLVQPTRLLEYLEEKYEHEHLYERDEGDKWRKKFELGLEFPNLPYYID 60
QY 61 GDVKLTQSMALIRYIADKHNMLGCGCPKERAETISMLEGAVLDIRYGVSRAYSXDFETLVK 120
DB 61 GDVKLTQSMALIRYIADKHNMLGCGCPKERAETISMLEGAVLDIRYGVSRAYSXDFETLVK 120
QY 121 DFLSKLPEMLKMFEDRLCKHTYINGDHVTHPPDMLYDALDVLVYMPMCLDAPFKLVCKR 180
DB 121 DFLSKLPEMLKMFEDRLCKHTYINGDHVTHPPDMLYDALDVLVYMPMCLDAPFKLVCKR 180
QY 181 KRLEAIPOIDKYLKSSKTYIAMPLOGQATFGGSDHPKSDLVPRGSPNSRVDS-----S 234

Db 181 KRIEAIPOIDKYLKSKYIAMPLOGWATFGGDHPKSDLVPRGSPRASYVHHHHH 240
QY 235 GRAETRIPIKYGKSLR 251
Db 241 HHSVTKMSYVVELLR 257

RESULT 8
PCT-US94-09700-6

Sequence 6, Application PC/TUS9409700
GENERAL INFORMATION:
APPLICANT: BAYLOR COLLEGE OF MEDICINE
APPLICANT: SMITH, JAMES R.
TITLE OF INVENTION: SENESCENT CELL DERIVED INHIBITORS OF
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOWREY & SIMON
STREET: 1299 PENNSYLVANIA AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/09700
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/808,523
FILING DATE: 16-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/970,462
FILING DATE: 02-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/113,372
FILING DATE: 30-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/153,564
FILING DATE: 17-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/203,535
FILING DATE: 25-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/229,420
FILING DATE: 15-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/274,535
FILING DATE: 13-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: AUERBACH, JEFFREY I.
REGISTRATION NUMBER: 32,680
REFERENCE/DOCKET NUMBER: 225-102-CIP7-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 383-7451
TELEFAX: (202) 383-6610
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 232 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Schistosoma japonicum
IMMEDIATE SOURCE:
CLONE: GST
PCT-US94-09700-6

Query Match 66.3%; Score 1227; DB 4; Length 232;
Best Local Similarity 98.3%; Pred. No. 2,6e-105;
Matches 228; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MSPILGYWKIKGLVQPTRLLEYLEEYEEHLVYERDEGDKWRNKKFELGLEFPNLPYYID 60
Db 1 MSPILGYWKIKGLVQPTRLLEYLEEYEEHLVYERDEGDKWRNKKFELGLEFPNLPYYID 60
QY 61 GDVKLTQSMALIRIYIAKHNMLGGCPKERAEISMLEGAVDIRGVSRIASKDPETLV 120
Db 61 GDVKLTQSMALIRIYIAKHNMLGGCPKERAEISMLEGAVDIRGVSRIASKDPETLV 120
QY 121 DFLSKLPEMLKMFEDRLCKTYLNGDHYTHDPFMYDALDVLYMDPCKDAFPKLVCFK 180
Db 121 DFLSKLPEMLKMFEDRLCKTYLNGDHYTHDPFMYDALDVLYMDPCKDAFPKLVCFK 180
QY 181 KRIEAIPOIDKYLKSKYIAMPLOGWATFGGDHPKSDLVPRGSPNSRVD 232
Db 181 KRIEAIPOIDKYLKSKYIAMPLOGWATFGGDHPKSDLVPRGSPNSRVD 232

RESULT 9
PCT-US94-09700-11

Sequence 11, Application PC/TUS9409700
GENERAL INFORMATION:
APPLICANT: BAYLOR COLLEGE OF MEDICINE
APPLICANT: SMITH, JAMES R.
TITLE OF INVENTION: SENESCENT CELL DERIVED INHIBITORS OF
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOWREY & SIMON
STREET: 1299 PENNSYLVANIA AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/09700
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/808,523
FILING DATE: 16-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/970,462
FILING DATE: 02-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/113,372
FILING DATE: 30-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/153,564
FILING DATE: 17-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/203,535
FILING DATE: 25-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/229,420
FILING DATE: 15-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/274,535
FILING DATE: 13-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: AUERBACH, JEFFREY I.
REGISTRATION NUMBER: 32,680
REFERENCE/DOCKET NUMBER: 225-102-CIP7-PCT

TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 383-7451
TELEFAX: (202) 383-6610
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 397 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Protein
HYPOTHETICAL: NO
IMMEDIATE SOURCE:
CLONE: GST-SDI-1 fusion protein
PCT-US94-09700-11

Query Match 66.3%; Score 1226; DB 4; Length 397;
Best Local Similarity 100.0%; Pred. No. 6,5e-105;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSPILGYWKIKGLVQPTRLLEYLEEKYEHLERDEGDKWRNKKFELGLEFPNLPYYID 60
Db 1 MSPILGYWKIKGLVQPTRLLEYLEEKYEHLERDEGDKWRNKKFELGLEFPNLPYYID 60
Qy 61 GGVKLTQSMATIRYIADKHNMLGCGCPKERAELISMLGAVLDIRYGSRIAYSKDFETLKY 120
Db 61 GGVKLTQSMATIRYIADKHNMLGCGCPKERAELISMLGAVLDIRYGSRIAYSKDFETLKY 120
Qy 121 DFLSKLPEMLKMFEDRLCHRTYLNCGDHVTHPDEMLYDALDVLYLMDPMCLDAFPKLVCFK 180
Db 121 DFLSKLPEMLKMFEDRLCHRTYLNCGDHVTHPDEMLYDALDVLYLMDPMCLDAFPKLVCFK 180
Qy 181 KRLEAIPOIDKYLKSSKYIAMPLOGWQATFGGSDHPKSDLVPRGSP 227
Db 181 KRLEAIPOIDKYLKSSKYIAMPLOGWQATFGGSDHPKSDLVPRGSP 227

RESULT 10
US-08-974-549A-605
Sequence 605, Application US/08974549A
Patent No. 6166178

GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-00261005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0300
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 605:
SEQUENCE CHARACTERISTICS:
LENGTH: 514 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..515
OTHER INFORMATION: /note="fusion protein composed of
glutathione-S-transferase sequence,
thrombin cleavage sequence, recognition
sequence for heart muscle protein
OTHER INFORMATION: kinase, residues introduced by cloning
OTHER INFORMATION: and hTERT protein fragment"

US-08-974-549A-605

Query Match 66.3%; Score 1226; DB 3; Length 514;
Best Local Similarity 96.6%; Pred. No. 9.1e-105;
Matches 229; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 MSPILGYWKIKGLVQPTRLLEYLEEKYEHLERDEGDKWRNKKFELGLEFPNLPYYID 60
Db 1 MSPILGYWKIKGLVQPTRLLEYLEEKYEHLERDEGDKWRNKKFELGLEFPNLPYYID 60
Qy 61 GGVKLTQSMATIRYIADKHNMLGCGCPKERAELISMLGAVLDIRYGSRIAYSKDFETLKY 120
Db 61 GGVKLTQSMATIRYIADKHNMLGCGCPKERAELISMLGAVLDIRYGSRIAYSKDFETLKY 120
Qy 121 DFLSKLPEMLKMFEDRLCHRTYLNCGDHVTHPDEMLYDALDVLYLMDPMCLDAFPKLVCFK 180
Db 121 DFLSKLPEMLKMFEDRLCHRTYLNCGDHVTHPDEMLYDALDVLYLMDPMCLDAFPKLVCFK 180
Qy 181 KRLEAIPOIDKYLKSSKYIAMPLOGWQATFGGSDHPKSDLVPRGSPNSVSSGRA 237
Db 181 KRLEAIPOIDKYLKSSKYIAMPLOGWQATFGGSDHPKSDLVPRGSPNSVSSGRA 237

RESULT 11
US-08-864-224-11
Sequence 11, Application US/08864224

```
Patent No. 5851808
GENERAL INFORMATION:
APPLICANT: Ellledge, Stephen J.
APPLICANT: Liu, Qinghua
TITLE OF INVENTION: Rapid subcloning Using Site-Specific
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESSES:
ADDRESS: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent'n Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/864,224
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: BCM-02681
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 579 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-864-224-11

Query Match      66.2%; Score 1224; DB 2; Length 579;
Best Local Similarity 97.9%; Pred. No. 1.6e-104;
Matches 228; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
```

```
ADDRESSEE: Pollock, Vande Sande & Priddy
STREET: 1990 M Street, NW, Suite 800
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent'n Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/395,507
FILING DATE: 27-FEB-1995
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP Hei-6-54672
FILING DATE: 28-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Amernick, Burton A.
REGISTRATION NUMBER: 24,852
REFERENCE/DOCKET NUMBER: 1581/007
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 331-7111
TELEFAX: (202) 293-6229
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 352 amino acids
TYPE: amino acid
STRADENESS:
TOPOLOGY: not relevant
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: rabbit
FEATURE:
NAME/KEY: Protein
LOCATION: 1..352
OTHER INFORMATION: /note="G15 Antigen"
US-08-395-507-1

Query Match      66.1%; Score 1222; DB 1; Length 352;
Best Local Similarity 83.0%; Pred. No. 1.3e-104;
Matches 234; Conservative 9; Mismatches 17; Indels 22; Gaps 3;
```

APPLICANT: Guemene, Daniel
APPLICANT: Zadowny, David
APPLICANT: Karatzos, Costas
TITLE OF INVENTION: USE OF PROLACTIN FOR PREVENTING AND/OR
TREATING BIRD BROODINESS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: WEISER & ASSOCIATES
STREET: 230 South Fifteenth Street, Suite 500
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19102
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/737,248
FILING DATE: 28-APR-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR95/00576
FILING DATE: 03-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 94/05550
FILING DATE: 05-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Weiser, Gerard J.
REGISTRATION NUMBER: 19,763
REFERENCE/DOCKET NUMBER: 989.6411P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-875-8383
TELEFAX: 215-875-8394
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 426 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-737-248-4

Query Match 65.9%; Score 1219.5; DB 3; Length 426;
Best Local Similarity 97.4%; Pred. No. 2.8e-104;
Matches 229; Conservative 2; Mismatches 3; Indels 1; Gaps 1;
QY 1 MSPILGWYKIKGLVOPTRLLLEYLEEKYEHLERDEGDKWRNKKFELGLEFNPPLYID 60
DB 1 MSPILGWYKIKGLVOPTRLLLEYLEEKYEHLERDEGDKWRNKKFELGLEFNPPLYID 60
QY 61 GGVKLTQSMALIRYIADKNHMLGCGPKERAELISMLGAVLDIRGYVSRIAYSDFTLKY 120
DB 61 GGVKLTQSMALIRYIADKNHMLGCGPKERAELISMLGAVLDIRGYVSRIAYSDFTLKY 120
QY 121 DFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDMLYDALDVLVLYMPMCDAPPKLVCFK 180
DB 121 DFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDMLYDALDVLVLYMPMCDAPPKLVCFK 180
QY 181 KRLEAIPOIDKYLKSSKYIAMPLOGMOATFGGDPKSDLVPRGS-SLPICSSG 235
DB 181 KRLEAIPOIDKYLKSSKYIAMPLOGMOATFGGDPKSDLVPRGS-SLPICSSG 234

RESULT 14
US-08-395-507-2
Sequence 2, Application US/08395507
Patent No. 5578456
GENERAL INFORMATION:
APPLICANT: Fujimura, Katsuya
APPLICANT: Ueno, Eiichi
APPLICANT: Fujii, No. 5578456uyuk1

APPLICANT: Okada, Masahisa
TITLE OF INVENTION: Anti-treponema Pallidum Antibody
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pollock, Vande Sande & Priddy
STREET: 190 W Street, NW, Suite 800
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/395,507
FILING DATE: 27-FEB-1995
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP Hei-6-54672
FILING DATE: 28-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Amerinick, Burton A.
REGISTRATION NUMBER: 24,852
REFERENCE/DOCKET NUMBER: 1581/007
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 331-7111
TELEFAX: (202) 293-6229
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 362 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: not relevant
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: Protein
LOCATION: 1..362
OTHER INFORMATION: /note="G17 Antigen"
US-08-395-507-2

Query Match 65.9%; Score 1219; DB 1; Length 362;
Best Local Similarity 100.0%; Pred. No. 2.5e-104;
Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSPILGWYKIKGLVOPTRLLLEYLEEKYEHLERDEGDKWRNKKFELGLEFNPPLYID 60
DB 1 MSPILGWYKIKGLVOPTRLLLEYLEEKYEHLERDEGDKWRNKKFELGLEFNPPLYID 60
QY 61 GGVKLTQSMALIRYIADKNHMLGCGPKERAELISMLGAVLDIRGYVSRIAYSDFTLKY 120
DB 61 GGVKLTQSMALIRYIADKNHMLGCGPKERAELISMLGAVLDIRGYVSRIAYSDFTLKY 120
QY 121 DFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDMLYDALDVLVLYMPMCDAPPKLVCFK 180
DB 121 DFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDMLYDALDVLVLYMPMCDAPPKLVCFK 180
QY 181 KRLEAIPOIDKYLKSSKYIAMPLOGMOATFGGDPKSDLVPRGS 226
DB 181 KRLEAIPOIDKYLKSSKYIAMPLOGMOATFGGDPKSDLVPRGS 226

RESULT 15
US-08-102-757-9
Sequence 9, Application US/08102757
Patent No. 5480864
GENERAL INFORMATION:
APPLICANT: Tajima, Masahiro

APPLICANT: Ohnuma, Manami
APPLICANT: Lerner, Ethan A.
TITLE OF INVENTION: Modified Maxadilan Protein, Its
TITLE OF INVENTION: Preparation and Use, and DNA encoding the Protein
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESS: Flanagan, Henderson, Farabow, Garrett &
ADDRESS: Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/102,757
FILING DATE: 01-JAN-1993
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 05136.0003-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 291 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-102-757-9

Query Match 65.2%; Score 1207; DB 1; Length 291;
Best Local Similarity 99.1%; Pred. No. 2,4e-103;
Matches 224; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MSPIIGYWKIKGLVQPTLLILEEKEYEHLVERDEGDKWRNKKFELGLEFPNLPYYTD 60
DB 1 MSPIIGYWKIKGLVQPTLLILEEKEYEHLVERDEGDKWRNKKFELGLEFPNLPYYTD 60
QY 61 GYVKLTQSMATIRYTADKNNMLGGCPKERAELSMLEGAVIDIRYGVSRITAYSKDPETLKV 120
DB 61 GYVKLTQSMATIRYTADKNNMLGGCPKERAELSMLEGAVIDIRYGVSRITAYSKDPETLKV 120
QY 121 DFLSKLPKMKFEDRLCHKTYLNGDVTNPDPMLYDALDVVLYMDPMCLDAFPKLVCFK 180
DB 121 DFLSKLPKMKFEDRLCHKTYLNGDVTNPDPMLYDALDVVLYMDPMCLDAFPKLVCFK 180
QY 181 KRLEAIPOIDKTYLKSKSYTAMPLOGWQATFGGDDHPPKSDLVPRGS 226
DB 181 KRLEAIPOIDKTYLKSKSYTAMPLOGWQATFGGDDHPPKSDLVPRGS 226

Search completed: March 18, 2001, 05:30:32
Job time: 3401 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2001, 04:58:12 ; Search time 66.39 Seconds
(without alignments)
350.805 Million cell updates/sec

Title: US-09-402-488a-2
Perfect score: 1850
Sequence: 1 MSPILGYWKIKGLVQPTRL.....PKQSHNDGDEEIPPEYLQ 343

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 6790655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1170	63.2	219 2	A26484 glutathione transf
2	990	53.5	218 2	A45556 glutathione S-tran
3	927	50.1	218 2	A45523 glutathione transf
4	654	35.4	209 2	A48388 glutathione S-tran
5	483.5	26.1	218 2	S33860 glutathione transf
6	483.5	26.0	218 2	A29794 glutathione transf
7	480.5	26.0	218 2	S32425 glutathione transf
8	480.5	26.0	218 2	A47486 glutathione transf
9	478.5	25.9	218 2	S65674 glutathione transf
10	474.5	25.6	218 2	A39375 glutathione transf
11	473.5	25.6	217 2	JX0095 glutathione transf
12	470.5	25.4	218 2	S13202 glutathione transf
13	469.5	25.4	218 2	A23732 glutathione transf
14	468.5	25.3	218 2	B34159 glutathione transf
15	467.5	25.3	218 2	S01719 glutathione transf
16	464.5	25.1	218 2	B29231 glutathione transf
17	464.5	25.1	218 2	A46048 glutathione transf
18	462.5	25.0	218 2	A429036 glutathione transf
19	456.5	24.7	218 1	XURG4 glutathione transf
20	453.5	24.5	218 2	A46143 mu-class glutathio
21	453.5	24.5	218 2	B28946 glutathione transf
22	449.5	24.3	225 2	A35295 glutathione transf
23	429.5	23.2	220 2	S18464 glutathione transf
24	385.5	20.8	219 2	S50146 major allergen Dpl
25	367	19.8	65 1	HULXH thrombin inhibitor
26	359	19.4	65 2	S78521 hirudin IIT - me
27	355	19.2	65 2	S05676 hirudin IIT - me
28	353	19.1	65 2	S78520 hirudin IITa - me
29	352	19.0	65 2	S05677 hirudin IITb - me

30	352	19.0	65 2	S05675 hirudin IITa - med
31	349	18.9	65 2	S05674 hirudin IITb - med
32	346	18.7	65 2	S05673 hirudin IITa - med
33	344	18.6	65 2	S05678 hirudin IIT - medic
34	339	18.3	65 2	S05679 hirudin IIT - medic
35	329	17.8	72 2	A37417 thrombin inhibitor
36	311	16.8	66 2	A24350 thrombin inhibitor
37	295	15.9	55 2	S05672 hirudin I - medic
38	257.5	13.9	142 2	S17462 glutathione transf
39	249	13.5	63 2	A53883 hirudin HVI homolo
40	248.5	13.4	208 2	S41933 glutathione transf
41	247	13.4	84 2	S33329 hirudin HM2 - leec
42	244	13.2	210 2	S71958 glutathione transf
43	242	13.1	203 2	S13780 glutathione transf
44	240	13.0	210 1	A37378 glutathione transf
45	238	12.9	210 2	S71957 glutathione transf

ALIGNMENTS

RESULT 1
A26484
glutathione transferase (EC 2.5.1.18) - fluke (Schistosoma japonicum) (fragment)
C:Species: Schistosoma japonicum
C:Date: 25-Oct-1987 #sequence_revision 30-Sep-1989 #text_change 03-Feb-1994
C:Accession: A94139; A26484; A28315
R:Smith, D.B.; Davern, K.M.; Board, P.G.; Tju, W.U.; Garcia, E.G.; Mitchell, G.F.
Proc. Natl. Acad. Sci. U.S.A. 83, 8703-8707, 1986
A:Title: Mr 26,000 antigen of Schistosoma japonicum recognized by resistant WEHI 129/
A:Reference number: A94139; MID:87041520
A:Accession: A94139
A:Molecule type: mRNA
A:Residues: 1-219 <SMI>
R:Smith, D.B.; Davern, K.M.; Board, P.G.; Tju, W.U.; Garcia, E.G.; Mitchell, G.F.
Proc. Natl. Acad. Sci. U.S.A. 84, 6541, 1987
A:Reference number: A94181
A:Contents: annotation; revision to residues 210-219
C:Superfamily: glutathione transferase
C:Keywords: transferase

Query Match 63.2% Score 1170; DB 2; Length 219;
Best Local Similarity 100.0%; Pred. No. 7.4e-89;
Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 PILGYWKIKGLVQPTRLLEYLEEKYEHLXERDEGKRNKKRELLEPPNLPYIDGD 62
DB 4 PILGYWKIKGLVQPTRLLEYLEEKYEHLXERDEGKRNKKRELLEPPNLPYIDGD 63
QY 63 VKLTQSMALIRYADKNMMLGCGCPKERAETSMLEGAVIDIRYGVSRATYASQDFETLKVDF 122
DB 64 VKLTQSMALIRYADKNMMLGCGCPKERAETSMLEGAVIDIRYGVSRATYASQDFETLKVDF 123
QY 123 LSKLPEMLKMFEDRLCKHTYTLNGSDVTHPPDFMLYDALDVLVYMDPKLDAFPKLVCEKKR 182
DB 124 LSKLPEMLKMFEDRLCKHTYTLNGSDVTHPPDFMLYDALDVLVYMDPKLDAFPKLVCEKKR 183
QY 183 IEALPQIDKYLKSSKRYTAMPLOQWATFFGGGDHPK 218
DB 184 IEALPQIDKYLKSSKRYTAMPLOQWATFFGGGDHPK 219
RESULT 2
A45556
glutathione S-transferase - fluke (Schistosoma mansoni)
C:Species: Schistosoma mansoni
C:Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 03-May-1996
C:Accession: A45556
R:Wright, M.D.; Harrison, R.A.; Melder, A.M.; Newport, G.R.; Mitchell, G.F.
Mol. Biochem. Parasitol. 49, 177-179, 1991
A:Title: Another 26-kilodalton glutathione S-transferase of Schistosoma mansoni.
A:Reference number: A45556; MID:92131046

A:Contents: clone pGT 875
 A:Accession: A28946
 A:Molecule type: mRNA
 A:Residues: 2-218 <PEA>
 A:Cross-references: GB:J03952; NID:9193687; P1DN:AAA37747.1; P1D:9309278
 R:Pearson, W.R.; Windle, J.J.; Morrow, J.F.; Benson, A.M.; Tatalay, P.
 J. Biol. Chem. 258, 2052-2062, 1983
 A:Title: Increased synthesis of glutathione S-transferases in response to anticarcinogen
 A:Reference number: A92411; MUID:83109018
 A:Accession: A20831
 A:Molecule type: protein
 A:Residues: 2-7, 'X', 8-13, 'X', 15-23, 'X', 25, 'XX', 28, 'N', 30-31, 'X', 33, 'X', 35-39, 'W', 41 <PEI
 R:Hayes, J.D.; Kerr, L.A.; Peacock, S.D.; Cronshaw, A.D.; Mclellan, L.I.
 Biochem. J. 277, 501-512, 1991
 A:Title: Hepatic glutathione S-transferases in mice fed on a diet containing the anticar
 cinogenic elution of the glutathione-S-transferase affinity matrix.
 A:Reference number: A24735; MUID:86042634
 A:Accession: 124735
 A:Molecule type: protein
 A:Residues: 2-7, 'X', 8-13, 'X', 15-23, 'X', 25 <MAN>
 R:Hayes, J.D.; Kerr, L.A.; Peacock, S.D.; Cronshaw, A.D.; Mclellan, L.I.
 Biochem. J. 277, 501-512, 1991
 A:Title: Hepatic glutathione S-transferases in mice fed on a diet containing the anticar
 cinogenic elution of the glutathione-S-transferase affinity matrix.
 A:Reference number: S11311; MUID:96268466
 A:Accession: S16933; MUID:91315425
 A:Molecule type: protein
 A:Residues: 2-36 <NAV>
 R:Fernandes, C.L.; Dong, J.H.; Roebuck, B.D.; Chisari, F.V.; Montali, J.A.; Schmidt Jr.,
 Arch. Biochem. Biophys. 331, 104-116, 1996
 A:Title: Elevations of hepatic quinine reductase, glutathione, and alpha- and mu-class g
 A:Reference number: S11311; MUID:96268466
 A:Accession: S11312
 A:Molecule type: protein
 A:Residues: 97-108 <FER>
 A:Gene: GSTM1
 A:Introns: 12/3; 38/1; 59/3; 87/1; 120/3; 152/3; 189/3
 C:Superfamily: glutathione transferase
 C:Keywords: dimer; transferase
 F:2-218/Product: glutathione transferase mul #status experimental <MAT>

Query Match 26.1%; Score 483.5; DB 2: Length 218;
 Best Local Similarity 44.2%; Pred. No. 1.8e-32;
 Matches 92; Conservative 38; Mismatches 73; Indels 5; Gaps 1;

QY 1 MSPILGKWKIKGIYOPTRLLEYLEEYEHLYERDEG-----DKRNKKFELGLEFPNL 55
 DB 1 MMLILGYWVWVRLTHIRILLETSSYDEKRYTMGDADFDPSQWLNKFKLGIDFEPNL 60
 QY 56 PYIIDGVKLTQSMALIRYIADKHNMLGCGPKRAEISMEGAVLDIRGVSRKATSKDF 115
 DB 61 PYLIDSSKRTITQSMALIRYIADKHNMLGCGPKRAEISMEGAVLDIRGVSRKATSKDF 120
 QY 116 EPLKADFLSKLEPMLEKMFEDRLCHKTLYLNGDHVTHPEMLYDALDVLYLMDPCGLDAFPK 175
 DB 121 EKQKEPFLKTIPEKMKLYSEFLGKRPWFAGDKRYTVDFLAYDILQYRMFEPKCLDAFPN 180
 QY 176 IWCFFKKRIAPQIDKYLKSSKYIAMPL 203
 DB 181 LQDFLAREFGLKRIKSYMKSSRYIATPI 208

RESULT 6
 A29794
 glutathione transferase (EC 2.5.1.18) class mu chain 3 - rat
 N:Alternate names: glutathione S-transferase Yb1; ligandin
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 28-Aug-1989 #sequence-revision 28-Aug-1989 #text-change 18-Jun-1999
 C:Accession: A29794; A25510; A24085; B61004; F24735; A33397; S27111; S17167; A26187
 R:Chang, C.; Saltzman, A.G.; Sorensen, N.S.; Hilpakka, R.A.; Liao, S.
 J. Biol. Chem. 262, 11901-11903, 1987
 A:Title: Identification of glutathione S-transferase Yb-1 mRNA as the androgen-repressed

A:Reference number: A29794; MUID:87308179
 A:Accession: A29794
 A:Molecule type: mRNA
 A:Residues: 1-218 <CH>
 A:Cross-references: GB:J02810; NID:9204514; P1DN:AAA1293.1; P1D:9204515
 R:Lai, H.C.J.; Grove, G.; Tu, C.P.D.
 Nucleic Acids Res. 14, 6101-6114, 1986
 A:Title: Cloning and sequence analysis of a cDNA for a rat liver glutathione S-transf
 A:Reference number: A25510; MUID:86312882
 A:Accession: A25510
 A:Molecule type: mRNA
 A:Residues: 1-218 <LAT>
 A:Cross-references: GB:X04229; NID:956337; P1DN:CAA27811.1; P1D:956338
 R:Ding, G.J.F.; Lu, A.Y.H.; Pickett, C.B.
 J. Biol. Chem. 260, 13268-13271, 1985
 A:Reference number: A24085; MUID:86033768
 A:Accession: A24085
 A:Molecule type: mRNA
 A:Residues: 1-198, 'NC', 201-218 <DIN>
 A:Cross-references: GB:M11719; NID:9204502; P1DN:AAA41287.1; P1D:9204503
 A:Experimental source: clone pGTA/C44
 R:Chang, L.H.; Hsieh, J.C.; Chen, W.L.; Tam, M.F.
 Electrophoresis 11, 589-593, 1990
 A:Title: Identification of rat liver glutathione S-transferase Yb subunits by partial
 lal isoelectric focusing gel.
 A:Reference number: A61004; MUID:91031411
 A:Accession: B61004
 A:Molecule type: protein
 A:Residues: 2-26 <CH2>
 R:Mannervik, B.; Alin, P.; Guthenberg, C.; Jonsson, H.; Tahir, M.K.; Warholm, M.; Jor
 Proc. Natl. Acad. Sci. U.S.A. 82, 7202-7206, 1985
 A:Title: Identification of three classes of cytosolic glutathione transferase common
 A:Reference number: A24735; MUID:86042634
 A:Accession: F24735
 A:Molecule type: protein
 A:Residues: 2-20 <MAN>
 R:Hsieh, J.C.; Liu, L.F.; Chen, W.L.; Tam, M.F.
 Biochem. Biophys. Res. Commun. 162, 1147-1154, 1989
 A:Title: Expression of Yb-1 glutathione S-transferase using a baculovirus expression
 A:Reference number: A33397; MUID:89350924
 A:Accession: A33397
 A:Molecule type: protein
 A:Residues: 2-21, 212-218 <HS1>
 R:Katuz, R.M.; Bono, B.; Colman, R.F.
 Arch. Biochem. Biophys. 298, 667-677, 1992
 A:Title: Identification of Tyr(115) labeled by S-(4-Dromo-2,3-dioxobutyl)glutathione
 A:Reference number: S27111; MUID:93037509
 A:Accession: S27111
 A:Molecule type: protein
 A:Residues: 2-20; 83-86, 'X', 88-96; 109-115, 'X', 117-122 <KAT>
 R:Hsieh, J.C.; Huang, S.C.; Chen, W.L.; Lai, Y.C.; Tam, M.F.
 Biochem. J. 278, 293-297, 1991
 A:Title: Cysteine-86 is not needed for the enzymic activity of glutathione S-transfer
 A:Reference number: S17167; MUID:91354218
 A:Accession: S17167
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 2-218 <HS2>
 C:Superfamily: glutathione transferase
 C:Keywords: dimer; transferase

Query Match 26.1%; Score 483.5; DB 2: Length 218;
 Best Local Similarity 44.2%; Pred. No. 1.8e-32;
 Matches 92; Conservative 38; Mismatches 73; Indels 5; Gaps 1;

QY 1 MSPILGKWKIKGIYOPTRLLEYLEEYEHLYERDEG-----DKRNKKFELGLEFPNL 55
 DB 1 MMLILGYWVWVRLTHIRILLETSSYDEKRYTMGDADFDPSQWLNKFKLGIDFEPNL 60
 QY 56 PYIIDGVKLTQSMALIRYIADKHNMLGCGPKRAEISMEGAVLDIRGVSRKATSKDF 115
 DB 61 PYLIDSSKRTITQSMALIRYIADKHNMLGCGPKRAEISMEGAVLDIRGVSRKATSKDF 120

R:Primiano, T.; Novak, R.F.
Arch. Biochem. Biophys. 301, 404-410, 1993
A:Title: Purification and characterization of class 'mu' glutathione S-transferase isozyme
A:Reference number: S30380; MUID:99213177
A:Accession: S30380
A:Molecule type: protein
A:Residues: 221 <PRD>
C:Superfamily: glutathione transferase
:Keywords: dimer; transferase

[illegible]

RESULT 10

A39375
glutathione transferase (EC 2.5.1.18) mu-2 [validated] - human

N:Alternate names: glutathione S-transferase GST4, muscle; glutathione transferase GST6,
C:Species: Homo sapiens (man)
C:Date: 20-Mar-1992 #sequence_revision 06-Jan-1995 #text_change 21-Jul-2000
C:Accession: A39375; A28810; C35187; S13305; S14344
R:Vorachek, W.R.; Pearson, W.R.; Rule, G.S.
Proc. Natl. Acad. Sci. U.S.A. 88, 4443-4447, 1991

A:Title: Cloning, expression, and characterization of a class-mu glutathione transferase
A:Reference number: A39375; MUID:91239584
A:Accession: A39375

A:Molecule type: mRNA
A:Residues: 1-218 <VOR>

A:Cross-references: GB:M63509; NID:g1833300; PIDN:AAA60963.1; PID:g1833301

A:Experimental source: myoblast

A:Note: translation of Initiator Met is not shown

R:Singh, S.V.; Ahmad, H.; Kurosky, A.; Awasthi, Y.C.
Arch. Biochem. Biophys. 264, 13-22, 1988

A:Title: Purification and characterization of unique glutathione S-transferases from human
A:Reference number: A90081; MUID:86280250
A:Accession: A28810

A:Molecule type: protein
A:Residues: 2-7, 'X',9-10, 'X',12-14 <SIN>
A:Experimental source: skeletal muscle

R:Tsuchida, S.; Maki, T.; Sato, K.
J. Biol. Chem. 265, 7150-7157, 1990

A:Title: Purification and characterization of glutathione transferases with an activity
A:Reference number: A35187; MUID:90237002
A:Accession: C35187

A:Molecule type: protein
A:Residues: 2-10, 'X',12-14, 'X',16-17, 'X',19-21, 'XXX' <TSU>
A:Experimental source: heart

R:Hussey, A.J.; Kerr, L.A.; Cronshaw, A.D.; Harrison, D.J.; Hayes, J.D.
Biochem. J. 273, 323-332, 1991

A:Title: Variation in the expression of Mu-class glutathione S-transferase isoenzymes from
A:Accession: S13305; MUID:91119559
A:Accession: S13305

A:Status: preliminary
A:Molecule type: protein
A:Residues: 2-25 <HUS>

A:Experimental source: skeletal muscle
R:Suzuki, T., Shaw, D.C., Board, P.G.
Biochem. J. 274, 405-408, 1991
A>Title: Purification and characterization of acidic glutathione S-transferase 6 from
A:Reference number: S14344; MUID:91174747
A:Accession: S14344
A:Molecule type: protein
A:Residues: 2-29, 'K', 31 <SUZ>
A:Experimental source: brain
C:Genetics:
A:Gene: GDB:GSTM2; GST4
A:Cross-references: GDB:128593; OMIM:138380
A:Map position: 1p13.3-1p13.3
C:Superfamily: glutathione transferase
C:Keywords: brain; cardiac muscle; heart; homodimer; skeletal muscle; transferase
P:2-18/Product: glutathione transferase mu-2 *status experimental <Mat>

Query Match	25.6%	Score 474.5	DB 2	Length 218
Best Local Similarity	44.0%	Pred. No. 1e-31		
Matches	95	Conservative	35	Mismatches 81
				Indels 5
				Gaps 1
QY	1	MSPILGWYKIGLVOPTRLLLEYLEEKYEHL	YERDEG----	DKWRNKKFELGIEFPNL 55
Db	1	MPMTGLWYINRILASIRLLEYETDSSEYEKKY	TGADADYDRSOWLNEKFEKIGIDPVL 60	
QY	56	PVYIDGVKYLTSMAIIRYIADKNHMLGGCCP	KERAEISMLCEAVLDIRGVSRIASKDF 115	
Db	61	PYLIDGTHKRTDSNMLRIARKNHMLCGESEK	EQREDLLEQFMDSRQMLKLCTDPDF 120	
QY	116	ETLKVDFLSKLPEMLKMEFDRLC	HKTYLNGDVTHTPDPVLVDALDVLVLMDCPMCLDAEPK 175	
Db	121	EKLKEVYLOALPEMLKLYSQFLGKQPMW	FGDKITFEVDIAYDVLERNQVFEPSCLDAPFN 180	
QY	176	IYCEKKRTAEIPIQIDKYLSSKYYIAMPI	DQMGATG 211	
Db	181	LKDPLSRFEGLEKISAYMKSSRFLPRP	VTFAVWG 216	

RESULT 11
JX0095
glutathione transferase (EC 2.5.1.18) b - guinea pig
N:Alternate names: glutathione S-alkyltransferase; glutathione S-aryltransferase; glutathione S-transferase
C:Species: *Cavia porcellus* (guinea pig)
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 18-Jun-1993
C:Accession: JX0095
R:Kamei, K.; Oshino, R.; Hara, S.
J. Biochem. 107, 111-117, 1990
A:Title: Amino acid sequence of glutathione S-transferase b from guinea pig liver.
A:Reference number: JX0095; MUID:90236961.
A:Accession: JX0095
A:Molecule type: protein
A:Residues: 1-217 <KAM>
C:Comment: Glutathione transferases are a family of enzymes that are multifunctional
C:Superfamily: glutathione transferase
C:Keywords: transferase

	Query Match Similarity	25.6%	Score 473.5	DB 2:	Length 217;
	' Best Local Similarity	45.1%	Pred. No. 1.2e-31;		
	Matches	92;	Conservative	34;	Mismatches 73; Indels 5; Gaps 1;
OY	5 LGYWKIKGLVQPTRLLEYLEEKYEENHLYERDEG-----DKWRNKKFELGTGLEFPNLPRYT	59			
Dd	4 LGYWNIRLGHTRIRLLLETNGSEEEKRRNMGDADVDYDRSQMLKFKIGDFFNPPLYI	63			
OY	60 DGDVKLTGSMALIRYATADKNHNLGCCPKERAETISMLEGAVALDIRGVSRIVAYSKDQETLK	119			
Dd	64 DETHLTGSNALIRYIAKRNHNLCGTEEEETTIRMDLLENVMQIMRIQLIMLCSPDEQRK	123			
OY	120 VDFLSKLREPMKFEDRCLHKTYLGNDINHTHDMLYALDALDVLTMDPNCDAFAFLVSCF	179			
Dd	124 ADFLLIPDPKMKEISQFGLKLPWEGNKTLTVDFLAIDVLDDRYMILERPCLCEAFANLADF	183			

OY 180 KKRIEAIPOIDKYKSSKYIAMPL 203
DB 184 ISREFGLEKISSYKSSRFLPKPL 207

RESULT 12

glutathione transferase (EC 2.5.1.18) Y1 - Chinese hamster
C:Species: Citicellus griseus (Chinese hamster)
C:Date: 18-Feb-1994 #sequence_revision 06-Jan-1995 #text_change 18-Jun-1999
C:Accession: S13202
R:Robert de Saint Vincent, B.; Hyrien, O.; Debatisse, M.; Buttin, G.
A:Title: Coamplification of mu class glutathione S-transferase genes and an adenylate de
A:Reference number: S13202; MUID:91031445
A:Accession: S13202
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-218 <ROB>
A:Cross-references: GB:X57489; NID:949541; PIDN:CAA0726.1; PID:949542
A:Note: In Genbank entry CUI1, release 109, the source is designated as Citicellus longi
C:Superfamily: glutathione transferase
C:Keywords: dimer; transferase

Query Match 25.4%; Score 470.5; DB 2; Length 218;
Best Local Similarity 43.3%; Pred. No. 2,2e-31;

Matches 90; Conservative 36; Mismatches 77; Indels 5; Gaps 1;

OY 1 MSPILGWMKIGLVPTRLLEYLEEKEYEHL.YERDEG-----DKWRNKKFELGLEFPNL 55
DB 1 MPMILGWMVWVGLFNPRLLEYLETDSYBEKKYTMGDADPSQWLNKFKLGDPEPL 60
OY 56 PYIDGDKVLQSMALIRIADKHNMLGGCPKERAISMLSEAVDIDIRGVRIAYSKEF 115
DB 61 PYLDGSHKITYOSNALIRYLARKHNLGCTEERIRIVDLVENQAMDTROLMLCYPDPF 120
OY 116 ETLKVDFLSKLPEMLKMFEDRLCHKTLYLNGDVTHPDPMLYDALDVLYMDPMDLADFPK 175
DB 121 EKQKEPEFLKTIPEKMKMYSEFLGKRPWFAGDKVITLCGFLAYDVOYQMEFEKCLDPFN 180
OY 176 LVCFKKRIEAIPOIDKYKSSKYIAMPL 203
DB 181 LKDFLAREGLKISSAYMKTSRFLRPPI 208

RESULT 13

A23732
glutathione transferase (EC 2.5.1.18) mu - golden hamster
C:Species: Mesocricetus auratus (golden hamster)
C:Date: 30-Dec-1991 #sequence_revision 30-Dec-1991 #text_change 18-Jun-1999
C:Accession: A23732; S21908
R:Norris, J.S.; Schwartz, D.A.; Macleod, S.L.; Fan, W.; O'Brien, T.J.; Harris, S.E.; Tri
Mol. Endocrinol. 5, 979-986, 1991
A:Title: Cloning of a mu-class glutathione S-transferase complementary DNA and character
A:Reference number: A23732; MUID:92049380
A:Accession: A23732
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-218 <NOR>
A:Cross-references: EMBL:X61033; NID:949638; PIDN:CAA43368.1; PID:949639
A:Note: Submitted to the EMBL Data Library, July 1991
C:Superfamily: glutathione transferase
C:Keywords: transferase

Query Match 25.4%; Score 469.5; DB 2; Length 218;
Best Local Similarity 44.3%; Pred. No. 2,6e-31;
Matches 90; Conservative 37; Mismatches 71; Indels 5; Gaps 1;

OY 1 MSPILGWMKIGLVPTRLLEYLEEKEYEHL.YERDEG-----DKWRNKKFELGLEFPNL 55
DB 181 LKDFLAREGLKISSAYMKTSRFLRPPI 208

DB 1 MPTVLGWDINGLAHAIIRLLETTDTSYEKKYTMGDAPNEDRSQWLNKFKLGDPEPL 60
OY 56 PYIDGDKVLQSMALIRIADKHNMLGGCPKERAISMLSEAVDIDIRGVRIAYSKEF 115
DB 61 PYLDGSHKITYOSNALIRYLARKHNLGCTEERIRIVDLVENQAMDTROLMLCYPDPF 120

OY 116 ETLKVDFLSKLPEMLKMFEDRLCHKTLYLNGDVTHPDPMLYDALDVLYMDPMDLADFPK 175
DB 121 EKQKEPEFLKTIPEKMKMYSEFLGKRPWFAGDKVITLCGFLAYDVOYQMEFEKCLDPFN 180
OY 176 LVCFKKRIEAIPOIDKYKSSKY 198
DB 181 LKDFLAREGLKISSAYMKTSRFLRPPI 208

RESULT 14

B34159
glutathione transferase (EC 2.5.1.18) mu2 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 11-Apr-1990 #sequence_revision 25-Sep-1992 #text_change 18-Jun-1999
C:Accession: B34159; S30369; S30370; S30372
R:Townsend, A.T.; Goldsmith, M.E.; Pickett, C.B.; Cowan, K.H.
J. Biol. Chem. 264, 21582-21590, 1989
A:Title: Isolation, characterization, and expression in Escherichia coli of two murin
A:Reference number: A34159; MUID:90094327
A:Accession: B34159
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-218 <TON>
A:Cross-references: GB:J04666; NID:9193549; PIDN:AAA37706.1; PID:9309261
R:Awasthi, S.; Singhal, S.S.; Srivastava, S.K.; Awasthi, Y.C.
Arch. Biochem. Biophys. 301, 143-150, 1993
A:Title: Purification and characterization of glutathione S-transferase of murine ova
A:Reference number: S30369; MUID:93183007
A:Accession: S30369
A:Molecule type: protein
A:Residues: 2-25 <AW1>
A:Note: this form was identified by PI 6.7
A:Accession: S30370
A:Molecule type: protein
A:Residues: 2-25 <AW2>
A:Note: this form was identified by PI 7.9
A:Accession: S30372
A:Molecule type: protein
A:Residues: 2-25 <AW3>
A:Note: this form was identified by PI 8.5
C:Superfamily: glutathione transferase
C:Keywords: transferase

Query Match 25.3%; Score 468.5; DB 2; Length 218;
Best Local Similarity 42.8%; Pred. No. 3,2e-31;
Matches 89; Conservative 41; Mismatches 73; Indels 5; Gaps 1;

OY 1 MSPILGWMKIGLVPTRLLEYLEEKEYEHL.YERDEG-----DKWRNKKFELGLEFPNL 55
DB 1 MPMILGWMVWVGLFNPRLLEYLETDSYBEKKYTMGDADPSQWLNKFKLGDPEPL 60
OY 56 PYIDGDKVLQSMALIRIADKHNMLGGCPKERAISMLSEAVDIDIRGVRIAYSKEF 115
DB 61 PYLDGSHKITYOSNALIRYLARKHNLGCTEERIRIVDLVENQAMDTROLMLCYPDPF 120
OY 116 ETLKVDFLSKLPEMLKMFEDRLCHKTLYLNGDVTHPDPMLYDALDVLYMDPMDLADFPK 175
DB 121 EKQKEPEFLKTIPEKMKMYSEFLGKRPWFAGDKVITLCGFLAYDVOYQMEFEKCLDPFN 180
OY 176 LVCFKKRIEAIPOIDKYKSSKYIAMPL 203
DB 181 LKDFLAREGLKISSAYMKTSRFLRPPI 208

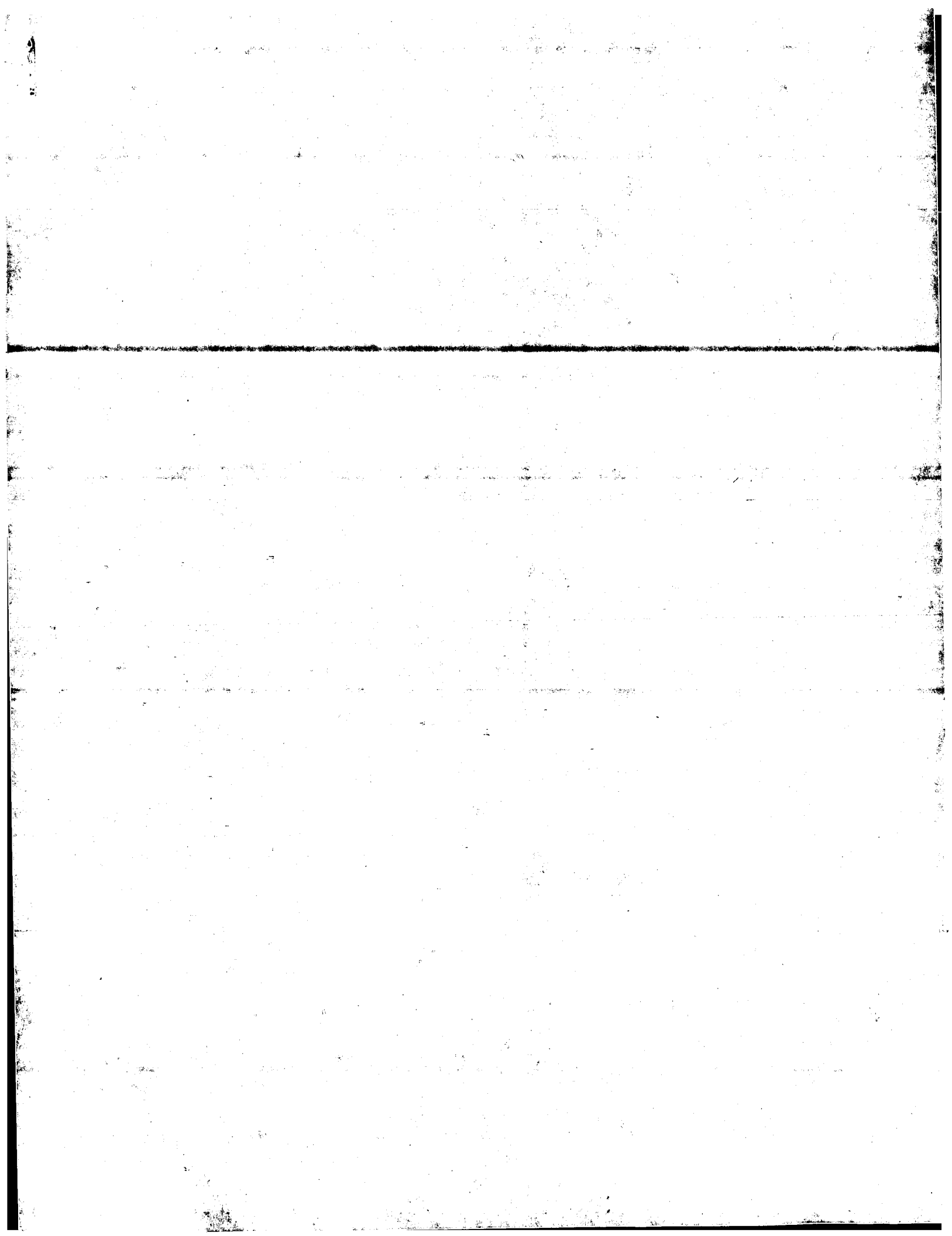
RESULT 15
S01719

glutathione transferase (EC 2.5.1.18) class mu, GSTM1 - human
 N:Alternate names: glutathione S-transferase 1, mu (H-b); glutathione transferase 4; glu
 C:Species: Homo sapiens (man)
 C:Date: 30-Jun-1989 #sequence, revision 30-Jun-1989 #ext, change 21-Jul-2000
 C:Accession: S01719, A30770; S32424; S13905; B22457; D35187; S27188; I51867; E24
 R:DeJong, J.L., Chang, C.M., Whang-Peng, J., Knutsen, T., Tu, C.P.D.
 Nucleic Acids Res. 16, 8541-8554, 1988
 A:Title: The human liver glutathione S-transferase gene superfamily: expression and chnc
 A:Reference number: S01719; MUID:88335606
 A:Contents: allele 1a (mu)
 A:Accession: S01719
 A:Molecule type: mRNA
 A:Residues: 1-218 <PE3>
 A:Cross-references: EMBL:X08020; NID:g31923; PIDN:CAA30821.1; PID:g31924
 A:Experimental source: liver
 R:Seidgard, J.; Vorachek, W.R.; Pero, R.W.; Pearson, W.R.
 Proc. Natl. Acad. Sci. U.S.A. 85, 7293-7297, 1988
 A:Title: Hereditary differences in the expression of the human glutathione transferase a
 A:Reference number: A30770; MUID:89017184
 A:Contents: allele 1b (psi)
 A:Accession: A30770
 A:Molecule type: mRNA
 A:Residues: 1-172, 'N', 174-218 <SEFI>
 A:Cross-references: EMBL:U03817; NID:g183668; PIDN:AAA59203.1; PID:g306812
 A:Experimental source: liver
 R:Zhong, S.; Spurr, N.K.; Hayes, J.D.; Wolf, C.R.
 Biochem. J. 291, 41-50, 1993
 A:Title: Deduced amino acid sequence, gene structure and chromosomal location of a novel
 A:Reference number: S32424; MUID:93228651
 A:Accession: S32424
 A:Molecule type: DNA
 A:Residues: 1-43, 'T', 45-172, 'N', 174-218 <ZHO>
 A:Cross-references: EMBL:X68676
 A:Note: the authors translated the codon ACG for residue 44 as Ser
 R:Comstock, K.E.; Sanderson, B.J.; Clafin, G.; Henner, W.D.
 Nucleic Acids Res. 18, 3670, 1990
 A:Title: GST1 gene deletion determined by polymerase chain reaction.
 A:Reference number: I37437; MUID:90301515
 A:Accession: I37437
 A:Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMH
 A:Molecule type: DNA
 A:Residues: 60-118 <RES>
 A:Cross-references: EMBL:X51451; NID:g31922; PIDN:CAA35817.1; PID:g4378985
 R:Singhal, S.S.; Ahmad, H.; Sharma, R.; Gupta, S.; Haque, A.K.; Awasthi, Y.C.
 Arch. Biochem. Biophys. 285, 64-73, 1991
 A:Title: Purification and characterization of human muscle glutathione S-transferases: e
 A:Reference number: S13905; MUID:91119426
 A:Accession: S13905
 A:Molecule type: protein
 A:Residues: 2-15 <SIN>
 A:Experimental source: muscle
 R:Alin, P.; Mannervik, B.; Jornvall, H.
 FEBS Lett. 182, 319-322, 1985
 A:Title: Structural evidence for three different types of glutathione transferase in hum
 A:Reference number: A91336; MUID:85154554
 A:Accession: B22457
 A:Molecule type: protein
 A:Residues: 2-24 <ALI>
 R:Tsuchida, S.; Maki, T.; Sato, K.
 J. Biol. Chem. 265, 7150-7157, 1990
 A:Title: Purification and characterization of glutathione transferases with an activit
 A:Reference number: A35187; MUID:90237002
 A:Accession: D35187
 A:Molecule type: protein
 A:Residues: 2-10, 'X', 12-14, 'X', 16-17, 'X', 19-25 <TSU>
 A:Experimental source: heart
 R:Singhal, S.S.; Saxena, M.; Awasthi, S.; Ahmad, H.; Sharma, R.; Awasthi, Y.C.
 Biochim. Biophys. Acta 1171, 19-26, 1992
 A:Title: Gender related differences in the expression and characteristics of glutathione
 A:Reference number: S27188; MUID:93042004
 A:Accession: S27188
 A:Molecule type: protein
 A:Residues: 2-13 <ST2>

A:Experimental source: colon
 R:Pearson, W.R.; Vorachek, W.R.; Xu, S.J.; Berger, R.; Hart, I.; Vannals, D.; Paters
 Am. J. Hum. Genet. 53, 220-233, 1993
 A:Title: Identification of class-mu glutathione transferase genes GSTM1-GSTM5 on huma
 A:Reference number: I51867; MUID:93304417
 A:Accession: I51867
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 125-172, 'N', 174-186 <RE2>
 A:Cross-references: GB:S62935; NID:g386429; PIDN:AA013938.1; PID:g4261638
 C:Comment: The GSTM1 locus for the mu isoenzyme of glutathione transferase is polymor
 stillbene oxide.
 C:Genetics:
 A:Gene: GDB:GSTM1; GSTM16
 A:Cross-references: GDB:120020; OMIM:138350
 A:Map position: 1p13.3-1p13.3
 A:Introns: 12/3; 38/1; 59/3; 87/1; 120/3; 152/3
 C:Superfamily: glutathione transferase
 C:Keywords: dimer; polymorphism; transferase
 F:2-218/Product: glutathione transferase class mu, GSTM1 #status predicted <MAT>

Query Match 25.3%; Score 467.5; DB 2; Length 218;
 Best Local Similarity 42.6%; Pred. No. 3, 8e-31;
 Matches 92; Conservative 39; Mismatches 80; Indels 5; Gaps 1;
 Oy 1 MSPILGYWKIKGLVQPRLLLEYLEEKYERHLYERDGE-----DKWENKKFELGLEFPNL 55
 Db 1 MPMLIGWMDIRGLAHAIIRLLLETDSSYEKKYTMGDAPDYDSQWLNENFKRGLDLPNL 60
 Oy 56 PYIIDGDVKLQSMALIRYIADKHNMLGCGPKERAETSMLEGAVALDIRYGVSRIVASKDF 115
 Db 61 PYLIDGAKHRTQSNALICYIARKNHNLGCEPTEEEKIRVDIENQTMDMHMLGIMCYNPEF 120
 Oy 116 ETLKVDPLSLPPELKKHFEERLCKITLNDYHTHPFMYLDALDVLVYNDPCKLDAFPK 175
 Db 121 EKLKPKYLELPKLPKLYSEFLKRPWFAGNKTITFVDVLVYDLHRIPEPCLDAPN 180
 Oy 176 LVCFKKRIEAIPOIDKYLKSSKYIAMPLOGMQATFG 211
 Db 181 LKDFISFEGLEKISATMKSSRFLPRPFVKMAVWG 216

Search completed: March 18, 2001, 05:33:19
 Job time: 2107 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2001, 05:30:37 ; Search time 71.87 Seconds

(without alignments)
154.124 Million cell updates/sec

Title: US-09-402-488a-2

Perfect score: 1850
Sequence: 1 MSPITGYWKIKGLVQPTRL.....PKQSHNDGDFEIEPEEYLQ 343

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the total score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	1179	63.7	218 1	GT26_SCHUA P08515 schistosoma
2	990	53.5	218 1	GT27_SCHUA P3561 schistosoma
3	927	50.1	218 1	GT26_SCHUA P15964 schistosoma
4	710.5	38.4	220 1	GT29_FASHE P56598 fasciola he
5	699	37.8	217 1	GT27_FASHE P31670 fasciola he
6	681	36.8	217 1	GT28_FASHE P31671 fasciola he
7	680	36.8	217 1	GT26_FASHE P30112 fasciola he
8	481.5	26.0	217 1	GTM1_MOUSE P10649 mus musculu
9	481.5	26.0	217 1	GTM1_MOUSE P04905 rattus norv
10	480.5	25.8	218 1	GTM4_HUMAN P03013 homo sapien
11	477.5	25.8	217 1	GTM4_HUMAN P46409 oryctolagus
12	473.5	25.6	217 1	GTM2_HUMAN P28161 homo sapien
13	473.5	25.6	217 1	GTM2_HUMAN P16413 cavia porce
14	468.5	25.3	217 1	GTM2_HUMAN P000285 cricetus
15	468.5	25.3	217 1	GTM2_HUMAN P30116 mesocricetu
16	467.5	25.3	217 1	GTM2_HUMAN P15626 mus musculu
17	465.5	25.2	217 1	GTM1_HUMAN P09488 homo sapien
18	463.5	25.1	217 1	GTM5_HUMAN P46439 homo sapien
19	461.5	24.9	217 1	GTM3_RAT P08009 rattus norv
20	455.5	24.6	217 1	GTM2_RAT P08010 rattus norv
21	452.5	24.5	217 1	GTM3_MOUSE P19630 mus musculu
22	449.5	24.3	225 1	GTM3_HUMAN P21266 homo sapien
23	437.5	23.6	224 1	GTM5_MOUSE P48774 mus musculu
24	432.5	23.4	219 1	GTM2_CHICK P20136 gallus gall
25	385.5	20.8	219 1	GTM1_DERPT P46419 dermatophag
26	367	19.8	65 1	ITH1_HIRME P01050 hirudo medi
27	359	19.4	65 1	ITHK_HIRME P28511 hirudo medi
28	355	19.2	65 1	ITHI_HIRME P28509 hirudo medi
29	353	19.1	65 1	ITHH_HIRME P28508 hirudo medi
30	352	19.0	65 1	ITHG_HIRME P28507 hirudo medi
31	352	19.0	65 1	ITHI_HIRME P28510 hirudo medi
32	349	18.9	65 1	ITHF_HIRME P28506 hirudo medi
33	346	18.7	65 1	ITHC_HIRME P28503 hirudo medi

34	344	18.6	65 1	ITHD_HIRME P28504 hirudo medi
35	339	18.3	65 1	ITHI_HIRME P28505 hirudo medi
36	329	17.8	72 1	ITH3_HIRME P09945 hirudo medi
37	311	16.8	66 1	ITH2_HIRME P09944 hirudo medi
38	300	16.2	65 1	ITHA_HIRME P28501 hirudo medi
39	251.5	13.6	208 1	GTB_DIRIM P46426 ditrofolaria
40	249	13.5	63 1	ITHV_HIRMA P81492 hirudinaia
41	248.5	13.4	208 1	GTB_ONCVO P46427 onchocerca
42	244	13.2	209 1	GTB_MESAU P06550 mesocricetu
43	242	13.1	207 1	GTB_PIG P80031 sus scrofa
44	240	13.0	209 1	GTB_HUMAN P09211 homo sapien
45	238	12.9	209 1	GTB_CRIMI P47954 cricetus

ALIGNMENTS

RESULT ID	GT26_SCHUA	STANDARD:	PRT:	218 AA.
AC	P08515:			
DT	01-AUG-1988 (rel. 08, Created)			
DT	30-AUG-1988 (rel. 08, Last sequence update)			
DT	30-MAY-2000 (rel. 39, Last annotation update)			
DE	GLUTATHIONE S-TRANSFERASE 26 KDA (EC 2.5.1.18) (GST 26) (S26 ANTIGEN)			
DE	(GST CLASS-ALPHA).			
OS	Schistosoma japonicum (Blood fluke).			
OC	Eukaryota; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea;			
OC	Strigolidae; Schistosomatidae; Schistosomatidae; Schistosoma.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=87041520; PubMed=3095841;			
RA	Smith D.B., Davern K.M., Board P.G., Tiu W.U., Garcia E.G.,			
RA	Mitchell G.F.;			
RA	Proc. Natl. Acad. Sci. U.S.A. 84:6541-6541(1987).			
RT	[3]			
RP	X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).			
RX	MEDLINE=95276631; PubMed=7538846;			
RA	Lim K., Ho J.X., Keeling K., Gilliland G.L., Ji X., Ruker F.,			
RA	Carter D.C.;			
RT	"Three-dimensional structure of Schistosoma japonicum glutathione S-			
RT	transferase fused with a six-amino acid conserved neutralizing			
RT	epitope of gp41 from HIV."			
RL	Protein Sci. 3:2233-2244(1994).			
RN	[4]			
RP	X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).			
RX	MEDLINE=95156484; PubMed=7853399;			
RA	McGill M., Williams D.R., Tainer J.A.;			
RT	"Crystal structures of a schistosomal drug and vaccine target:			
RT	glutathione S-transferase from Schistosoma japonica and its complex			
RT	with the leading antischistosomal drug praziquantel."			
RL	J. Mol. Biol. 246:21-27(1995).			
CC	- FUNCTION: CONUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER			
CC	OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.			
CC	- FUNCTION: GST ISOENZYMES APPEAR TO PLAY A CENTRAL ROLE IN THE			
CC	PARASITE DETOXIFICATION SYSTEM. OTHER FUNCTIONS ARE ALSO			
CC	SUSPECTED INCLUDING A ROLE IN INCREASING THE SOLUBILITY OF			
CC	HAEMATIN IN THE PARASITE GUT.			
CC	- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GUTATHIONE.			
CC	- SUBUNIT: HOMODIMER.			
CC	- MISCELLANEOUS: THERE ARE AT LEAST TWO ISOENZYMES OF GST IN			
CC	S. JAPONICUM.			
CC	- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. ALPHA FAMILY.			
CC	- This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: M14654; AAB55203.1; -
DR PIR: A26484; A26484.
DR PDB: 1GNE; 30-NOV-54.
DR PDB: 1GTA; 07-FEB-55.
DR PDB: 1GTB; 01-DEC-55.
DR PDB: 1B8X; 12-APR-59.
DR INTERPRO: IPR000521; -
DR PFM: PF00043; GST; 1.
DR Transferase; Antigen; Multigene family; 3D-structure.
KW SEQUENCE 218 AA; 25498 MW; 5E2AC418BD0EF13F CRC64;
SQ

Query Match 63.7%; Score 1179; DB 1; Length 218;
Best Local Similarity 100.0%; Pred. No. 3.8e-90;
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSPILGYMKIKGLVPTLLLEYLEKYEHLIERDEGDKMKKFEGLPEPNLPYYID 60
DB 1 MSPILGYMKIKGLVPTLLLEYLEKYEHLIERDEGDKMKKFEGLPEPNLPYYID 60
QY 61 GVKLTQSMALIRYIADKHNMLGCGPKERAETSMLEGAVIDIRYGVSRIVASKDEFTLV 120
DB 61 GVKLTQSMALIRYIADKHNMLGCGPKERAETSMLEGAVIDIRYGVSRIVASKDEFTLV 120
QY 121 DFLSKLPKMKFEEDRLCHKTYLNGDHYTHDPFMYDALDVLYMDPCLDAFPKLVCFK 180
DB 121 DFLSKLPKMKFEEDRLCHKTYLNGDHYTHDPFMYDALDVLYMDPCLDAFPKLVCFK 180
QY 181 KRIEAIPOIDKTKLSSKTYIAMPLOGMOTFFGGDHPK 218
DB 181 KRIEAIPOIDKTKLSSKTYIAMPLOGMOTFFGGDHPK 218

RESULT: 2

GT27_SCHMA STANDARD: PRT: 218 AA.

AC P35661.
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE GLUTATHIONE S-TRANSFERASE 26 KDA (EC 2.5.1.18) (GST 26) (SM26/2
DE ANTIGEN) (GST CLASS-ALPHA).
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea;
OC Strigoida; Schistosomatidae; Schistosomatidae; Schistosoma.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92131046; Pubmed=1775156;
RA Wright M.D., Harrison R.A., Weider A.M., Newport G.R., Mitchell G.F.;
RT "Another 26-kilodalton glutathione S-transferase of Schistosoma
mansoni".
RL Mol. Biochem. Parasitol. 49:177-179(1991).
CC -1- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
CC OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
CC -1- FUNCTION: GST ISOENZYMES APPEAR TO PLAY A CENTRAL ROLE IN THE
CC PARASITE DETOXIFICATION SYSTEM. OTHER FUNCTIONS ARE ALSO
CC SUSPECTED INCLUDING A ROLE IN INCREASING THE SOLUBILITY OF
CC HAEMATIN IN THE PARASITE GUT.
CC -1- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GLUTATHIONE.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. ALPHA FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: M73624; -; NOT_ANNOTATED_CDS.
DR PIR: A45556; A45556.
DR HSSP: P08515; 1GTB.
DR INTERPRO: IPR000521; -
DR PFM: PF00043; GST; 1.
DR Transferase; Antigen; Multigene family.
KW SEQUENCE 218 AA; 25411 MW; DBD3EE9028B36185 CRC64;
SQ

Query Match 53.5%; Score 990; DB 1; Length 218;
Best Local Similarity 82.6%; Pred. No. 1.4e-74;
Matches 180; Conservative 19; Mismatches 19; Indels 0; Gaps 0;

QY 1 MSPILGYMKIKGLVPTLLLEYLEKYEHLIERDEGDKMKKFEGLPEPNLPYYID 60
DB 1 MSPILGYMKIKGLVPTLLLEYLEKYEHLIERDEGDKMKKFEGLPEPNLPYYID 60
QY 61 GVKLTQSMALIRYIADKHNMLGCGPKERAETSMLEGAVIDIRYGVSRIVASKDEFTLV 120
DB 61 GVKLTQSMALIRYIADKHNMLGCGPKERAETSMLEGAVIDIRYGVSRIVASKDEFTLV 120
QY 121 DFLSKLPKMKFEEDRLCHKTYLNGDHYTHDPFMYDALDVLYMDPCLDAFPKLVCFK 180
DB 121 DFLSKLPKMKFEEDRLCHKTYLNGDHYTHDPFMYDALDVLYMDPCLDAFPKLVCFK 180
QY 181 KRIEAIPOIDKTKLSSKTYIAMPLOGMOTFFGGDHPK 218
DB 181 KRIEAIPOIDKTKLSSKTYIAMPLOGMOTFFGGDHPK 218

RESULT: 3

GT26_SCHMA STANDARD: PRT: 218 AA.

AC P15964.
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE GLUTATHIONE S-TRANSFERASE 26 KDA (EC 2.5.1.18) (GST 26) (SM26/1
DE ANTIGEN) (GST CLASS-ALPHA).
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea;
OC Strigoida; Schistosomatidae; Schistosomatidae; Schistosoma.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=PUERTO RICAN;
RX MEDLINE=90348716; Pubmed=2385266;
RA Trottein F., Kieny M.P., Verweide C., Torpier G., Pierce R.J.,
RA Balloul J.-M., Schmitt D., Lecocq J.-P., Capron A.;
RT "Molecular cloning and tissue distribution of a 26-kilodalton
RT Schistosoma mansoni glutathione S-transferase".
RL Mol. Biochem. Parasitol. 41:35-44(1990).
RN [2]
RP SEQUENCE OF 8-218 FROM N.A.
RX STRAIN=PUERTO RICAN;
RX MEDLINE=90271935; Pubmed=1693415;
RA Henkle K.J., Davern K.W., Wright M.D., Ramos A.J., Mitchell G.F.;
RT "Comparison of the cloned genes of the 26- and 28-kilodalton
RT glutathione S-transferases of Schistosoma japonicum and Schistosoma
mansoni".
RL Mol. Biochem. Parasitol. 40:23-34(1990).
CC -1- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
CC OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
CC -1- FUNCTION: GST ISOENZYMES APPEAR TO PLAY A CENTRAL ROLE IN THE
CC PARASITE DETOXIFICATION SYSTEM. OTHER FUNCTIONS ARE ALSO
CC SUSPECTED INCLUDING A ROLE IN INCREASING THE SOLUBILITY OF
CC HAEMATIN IN THE PARASITE GUT.
CC -1- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GLUTATHIONE.
CC -1- SUBUNIT: HOMODIMER.
CC -1- TISSUE SPECIFICITY: TEGUMENT AND IN SUBSEGMENTARY PARENCHYMAL
CC CELLS. GST 26 MAY BE ACTIVELY EXCRETED BY ADULT WORMS.

CC -1- MISCELLANEOUS: THERE ARE AT LEAST THREE ISOENZYMES OF GST IN
CC S.MANSONI.
CC -1- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. ALPHA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: M31106; AAA29888.1; -
CC EMBL: M26913; AAA29889.1; -
CC PIR: A45523; A45523.
CC HSSP: P08515; 1GTH.
CC INTERPRO: IPR000521; -
CC PFAM: PF00043; GST; 1.
CC Transferrase; Antigen; Multigene family.
CC SEQUENCE 218 AA; 25401 MW; 061A6548A842D6E8 CRC64;
KM
SQ

Query Match 50.1%; Score 927; DB 1; Length 218;
Best Local Similarity 79.8%; Pred. No. 2,1e-69;
Matches 174; Conservative 15; Mismatches 29; Indels 0; Gaps 0;

OY 1 MSPILGWKIKGLVPTRLLEYLEEKYEHLERDEGDMRNKKFELGLEFPNLPYYID 60
DB 1 MAPKGYMKVKGLVPTRLLEHLEETFERAYDRNEIDAMNDKFKLEFPNLPYYID 60
OY 61 GDUKLTQSMALIRYIADKHNMLGCPKRAEISMLEGAVLDIRYGVSRAYSDEFTLKY 120
DB 61 GDFKLTQSMALIRYIADKHNMLGACPKRAEISMLEGAVLDIRYGVSRAYSDEFTLKY 120
OY 121 DELSLPEMLKMFEDRLCKTYLNGDHYTHRPDMLYDALDYVLYMDPCLDAFPKLYCKR 180
DB 121 DFLNKLPGKLMKFEPLRSKRYTLNGCVTHPDMLYDALDYVLYMDPCLDAFPKLYCKR 180
OY 181 KRIFAIPOIDKYLKSSKRYAMPLOGMOATFGGSDHPK 218
DB 181 KCIEDLPQIKNTLNSSRYIKWPLQGMADATFGGSDHPK 218

RESULT 4
GT29_FASHE STANDARD: PRT; 220 AA.
AC P56598;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GLUTATHIONE S-TRANSFERASE 26 KDA 1 (EC 2.5.1.18) (GST1) (FHL) (GST
DE CLASS-ALPHA).
OS Fasciola hepatica (Liver fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea;
OC Echinostomida; Echinostomata; Fascioloidae; Fasciolidae; Fasciola.
RN [1]
RP MEDLINE=92155306; PubMed=1740183;
RX Panaccio M., Wilson L.R., Crameri S.L., Wjffels G.L., Spithill T.W.;
RT "Molecular characterization of cDNA sequences encoding glutathione S-
RN transiferases of Fasciola hepatica.";
RL Exp. Parasitol. 74:232-237(1992).
RN [2]
RP MEDLINE=94039664; PubMed=8224094;
RX Panaccio M., Wilson L.R., Crameri S.L., Wjffels G.L., Spithill T.W.;
RN Exp. Parasitol. 77:385-385(1993).
RN [3]
RP SEQUENCE OF 22-220 FROM N.A.
RA Crameri S.;
RA Patent number WO9008819, 09-AUG-1990.
CC -1- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
CC OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.

CC -1- FUNCTION: GST ISOENZYMES APPEAR TO PLAY A CENTRAL ROLE IN THE
CC PARASITE DETOXIFICATION SYSTEM. OTHER FUNCTIONS ARE ALSO
CC SUSPECTED INCLUDING A ROLE IN INCREASING THE SOLUBILITY OF
CC HAEMATIN IN THE PARASITE GUT.
CC -1- CATALYTIC ACTIVITY: RX + GLUTATHIONE -> HX + R-S-GUTATHIONE.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. ALPHA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: A00993; CA00118.1; -
CC INTERPRO: IPR000521; -
CC PFAM: PF00043; GST; 1.
CC Transferrase; Antigen; Multigene family.
CC SEQUENCE 220 AA; 25598 MW; 27B9F150B75D101F CRC64;
KM
SQ

Query Match 38.4%; Score 710.5; DB 1; Length 220;
Best Local Similarity 59.7%; Pred. No. 1.4e-51;
Matches 129; Conservative 32; Mismatches 54; Indels 1; Gaps 1;

OY 5 LGYWKIKGLVPTRLLEYLEEKYEHLERDEGDMRNKKFELGLEFPNLPYYIDGYK 64
DB 4 LGYWKIKGLVPTRLLEYLEEKYEHLERDEGDMRNKKFELGLEFPNLPYYIDGYK 62
OY 65 LTQSMALIRYIADKHNMLGCPKRAEISMLEGAVLDIRYGVSRAYSDEFTLKYDPLS 124
DB 65 LTQSMALIRYIADKHNMLGCPKRAEISMLEGAVLDIRYGVSRAYSDEFTLKYDPLS 122
OY 125 KLEPMKMFEDRLCKTYLNGDHYTHRPDMLYDALDYVLYMDPCLDAFPKLYCKR 184
DB 125 KLEPMKMFEDRLCKTYLNGDHYTHRPDMLYDALDYVLYMDPCLDAFPKLYCKR 182
OY 185 ALPOIDKYLKSSKRYAMPLOGMOATFGGSDHPKSD 220
DB 185 ALPSIKAYMESNRFITKWPNGHMAOFGGSDHPKSD 218

RESULT 5
GT27_FASHE STANDARD: PRT; 217 AA.
ID GT27_FASHE
AC P31670;
DT 01-JUL-1993 (Rel. 26, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GLUTATHIONE S-TRANSFERASE 26 KDA 47 (EC 2.5.1.18) (GST47) (FH47)
DE (GST CLASS-ALPHA).
OS Fasciola hepatica (Liver fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea;
OC Echinostomida; Echinostomata; Fascioloidae; Fasciolidae; Fasciola.
RN [1]
RP MEDLINE=92155306; PubMed=1740183;
RX Panaccio M., Wilson L.R., Crameri S.L., Wjffels G.L., Spithill T.W.;
RT "Molecular characterization of cDNA sequences encoding glutathione S-
RN transiferases of Fasciola hepatica.";
RL Exp. Parasitol. 74:232-237(1992).
RN [2]
RP MEDLINE=94039664; PubMed=8224094;
RX Panaccio M., Wilson L.R., Crameri S.L., Wjffels G.L., Spithill T.W.;
RN Exp. Parasitol. 77:385-385(1993).
RN [3]
RP SEQUENCE OF 22-220 FROM N.A.
RA Crameri S.;
RA Patent number WO9008819, 09-AUG-1990.
CC -1- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
CC OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.

RP SEQUENCE OF 7-105 FROM N.A.
 RA Cramerl S.;
 RL Patent number WO9008819, 09-AUG-1990.
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
 RX MEDLINE=98035725; PubMed=9367777;
 RA Rossjohn J., Fell S.C., Wilce M.C.J., Sexton J.L., Spithill T.W.,
 RA Parker M.W.;
 RT "Crystallization, structural determination and analysis of a novel
 RT parasite vaccine candidate: Fasciola hepatica glutathione
 RT S-transferase.";
 RL J. Mol. Biol. 273:857-872(1997).
 CC -1- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
 CC OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
 CC -1- FUNCTION: GST ISOENZYMES APPEAR TO PLAY A CENTRAL ROLE IN THE
 CC PARASITE DETOXIFICATION SYSTEM. OTHER FUNCTIONS ARE ALSO
 CC SUSPECTED INCLUDING A ROLE IN INCREASING THE SOLUBILITY OF
 CC HAEMATIN IN THE PARASITE GUT.
 CC -1- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GUTATHIONE.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. ALPHA FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M77681; AAA29140.1; -
 DR EMBL: A00996; CAA00121.1; -
 DR PDB: 1FHE; 29-JUL-98.
 DR INTERPRO: IPR000521; -
 DR PFM: PF00043; GST; 1.
 KW Transferase; Antigen; Multigene family; 3D-structure.
 FT INT_MET 0
 FT CONFLICT 65 65 T->I (IN REF. 3).
 FT CONFLICT 102 105 RIGF->FEEL (IN REF. 3).
 SQ SEQUENCE 217 AA; 25281 MW; 0FB8BFE63029E03 CRC64;
 Query Match 37.8%; Score 699; DB 1; Length 217;
 Best Local Similarity 58.2%; Pred. No. 1.2e-50;
 Matches 124; Conservative 36; Mismatches 53; Indels 0; Gaps 0;
 QY 5 LGYWKIRGLVPTRLLEYLEEKTEHLYERDEGDKMKNKFEELGEPNLPYYIDGVK 64
 DB 4 LGYWKIRGLVPTRLLEYLEEKTEHLYERDEGDKMKNKFEELGEPNLPYYIDGVK 63
 QY 65 LTOSMAITRYIADKHNMLGGCPKERAETISMEGAVLDIRGVSAIYASKDEFTLKVDFLS 124
 DB 64 LTOSVAIMRYIADKHNMLGGCPKERAETISMEGAVLDIRGVSAIYASKDEFTLKVDFLS 123
 QY 125 KLPEMLKMFEDRLCHKTYLNGDHYTHDPFMYLDALDVLYMDPCLDAFPKLVCFKKRIE 184
 DB 124 ELPLTKMWSDFLGDRHLDTSSVSHVDFMYETLDSTRYLAPCLDDEFPKLVCFKKRIE 183
 QY 185 AIPQIDKYLKSSKIYAMPLOGQWATFGGCHPP 217
 DB 184 ALPKIKAYMESKRIKWPFLNMAASFGADAP 216
 RESULT 6
 GT28_FASHE STANDARD; PRT; 217 AA.
 AC P31671;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE GLUTATHIONE S-TRANSFERASE 26 KDA 7 (EC 2.5.1.18) (GST7) (FH7) (GST
 CLASS-ALPHA).

OS Fasciola hepatica (liver fluke).
 OC Eukaryota; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea;
 OC Echinostomida; Echinostomata; Fasciolidae; Fasciolidae; Fasciola.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92155306; PubMed=1740183;
 RA Panaccio M., Wilson L.R., Cramerl S.L., Wjiffels G.L., Spithill T.W.,
 RA "Molecular characterization of cDNA sequences encoding glutathione S-
 RT transferases of Fasciola hepatica,"
 RL Exp. Parasitol. 74:232-237(1992).
 RN [2]
 RP ERRATUM.
 RX MEDLINE=94039664; PubMed=8224094;
 RA Panaccio M., Wilson L.R., Cramerl S.L., Wjiffels G.L., Spithill T.W.,
 RL Exp. Parasitol. 77:385-385(1993).
 RN [3]
 RP SEQUENCE OF 8-217 FROM N.A.
 RA Cramerl S.;
 RL Patent number WO9008819, 09-AUG-1990.
 CC -1- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
 CC OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
 CC -1- FUNCTION: GST ISOENZYMES APPEAR TO PLAY A CENTRAL ROLE IN THE
 CC PARASITE DETOXIFICATION SYSTEM. OTHER FUNCTIONS ARE ALSO
 CC SUSPECTED INCLUDING A ROLE IN INCREASING THE SOLUBILITY OF
 CC HAEMATIN IN THE PARASITE GUT.
 CC -1- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GUTATHIONE.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. ALPHA FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M77680; AAA29139.1; -
 DR EMBL: A00994; CAA00119.1; -
 DR HSSP: P08515; 1GTR.
 DR INTERPRO: IPR000521; -
 DR PFM: PF00043; GST; 1.
 KW Transferase; Antigen; Multigene family.
 FT INT_MET 0
 FT CONFLICT 35 42 NDERKMG->MIGRNGA (IN REF. 3).
 FT CONFLICT 188 196 IKRYKMSER->SRVRSRA (IN REF. 3).
 SQ SEQUENCE 217 AA; 25196 MW; 0099E1F59E49A9E CRC64;
 Query Match 36.8%; Score 681; DB 1; Length 217;
 Best Local Similarity 58.7%; Pred. No. 3.7e-49;
 Matches 125; Conservative 29; Mismatches 59; Indels 0; Gaps 0;
 QY 5 LGYWKIRGLVPTRLLEYLEEKTEHLYERDEGDKMKNKFEELGEPNLPYYIDGVK 64
 DB 4 LGYWKIRGLVPTRLLEYLEEKTEHLYERDEGDKMKNKFEELGEPNLPYYIDGVK 63
 QY 65 LTOSMAITRYIADKHNMLGGCPKERAETISMEGAVLDIRGVSAIYASKDEFTLKVDFLS 124
 DB 64 LTOSVAIMRYIADKHNMLGGCPKERAETISMEGAVLDIRGVSAIYASKDEFTLKVDFLS 123
 QY 125 KLPEMLKMFEDRLCHKTYLNGDHYTHDPFMYLDALDVLYMDPCLDAFPKLVCFKKRIE 184
 DB 124 GLPTTLKMSWDFLGDRYLIGSSVSHVDFMYETLDSTRYLAPCLDDEFPKLVCFKKRIE 183
 QY 185 AIPQIDKYLKSSKIYAMPLOGQWATFGGCHPP 217
 DB 184 DLPKIKYKMSERIKWPLHSHWTSFEGGDAP 216
 RESULT 7
 GT26_FASHE STANDARD; PRT; 217 AA.
 ID GT26_FASHE

AC P30112: 01-APR-1993 (rel. 25, Created)
 DT 15-DEC-1998 (rel. 37, Last sequence update)
 DT 15-DEC-1998 (rel. 37, Last annotation update)
 DE GLUTATHIONE S-TRANSFERASE 26 KDA 51 (EC 2.5.1.18) (GSM51) (FHS1)
 DE (GST CLASS-ALPHA).
 OS Fasciola hepatica (Liver fluke).
 OC Eukaryota; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea;
 OC Echinostomida; Echinostomata; Fascioloidae; Fasciolidae; Fasciola.
 RN 11
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92155306; PubMed=1740183;
 RA Panaccio M., Wilson L.R., Cramer S.L., Wiffels G.L., Spithill T.W.;
 RT "Molecular characterization of cDNA sequences encoding glutathione S-
 RT transferases of Fasciola hepatica."
 RL Exp. Parasitol. 74:232-237(1992).
 RN 12
 RP ERRATUM.
 RX MEDLINE=94039664; PubMed=8224094;
 RA Panaccio M., Wilson L.R., Cramer S.L., Wiffels G.L., Spithill T.W.;
 RL Exp. Parasitol. 77:385-385(1993).
 RN 13
 RP SEQUENCE OF 9-217 FROM N.A.
 RX MEDLINE=93228188; PubMed=7682383;
 RA Muro A., Rodriguez-Molina J.R., Hallyer G.V.;
 RT "Sequence analysis of a Fasciola hepatica glutathione S-transferase,
 RT cDNA clone."
 RL Am. J. Trop. Med. Hyg. 48:457-463(1993).
 CC -1- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
 CC OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
 CC -1- FUNCTION: GST ISOENZYMES APPEAR TO PLAY A CENTRAL ROLE IN THE
 CC PARASITE DETOXIFICATION SYSTEM. OTHER FUNCTIONS ARE ALSO
 CC SUSPECTED INCLUDING A ROLE IN INCREASING THE SOLUBILITY OF
 CC HAEMATIN IN THE PARASITE GUT.
 CC -1- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GUTATHIONE.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. ALPHA FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: M77682; AAA29141.1; -
 DR HSSP: P08515; 1GNE.
 DR INTERPRO: IPR000521; -
 DR PFAM: PF00043; GST_1.
 KW Transferase; Antigen; Multigene family.
 FT INIT MET 0 BY SIMILARITY.
 FT CONFLICT 83 83 T->S (IN REF. 3).
 FT CONFLICT 134 134 N->D (IN REF. 3).
 FT CONFLICT 146 146 P->T (IN REF. 3).
 SO SEQUENCE 217 AA; 25242 MW; EFCABFF7454EDC26 CRC64;

Query Match 36.8%; Score 680; DB 1; Length 217;
 Best Local Similarity 58.7%; Pred. No. 4,4e-49;
 Matches 125; Conservative 32; Mismatches 56; Indels 0; Gaps 0;

OY 5 LGYWKRGVOPTRLELEKEEHLVERDEGDMRNKKFELGJEPPLPYIDGDK 64
 DB 4 LGYWKRGVOPTRLELEKEEHLVERDEGDMRNKKFELGJEPPLPYIDGDK 63
 OY 65 LTQSMALIRYADKNNMLGCKPERAEISMLEGAVLDIRGVSRIVSKDFETLKVDLS 124
 DB 64 LTQSMALIRYADKNNMLGCKPERAEISMLEGAVLDIRGVSRIVSKDFETLKVDLS 123
 OY 125 KLPENLAKMEDRLCHKTYTHPDEMLYALADVLYLMDPKCLDAFFKLYCFKKRIE 184

DB 124 ELPTLKMNSNFIQDRHYLTGSPVSHVDFMVAELDCIRLAPQCLEDPFKLKEFSRIE 183
 OY 185 AIPQIDKYLKSSKYIAMPLQGWQATFGGDDHP 217
 DB 184 DLPIKAYMESSEKFIKWLNSWIASFCGGDAP 216

RESULT 8
 ID GTM1_MOUSE STANDARD; PRT; 217 AA.
 AC P10649;
 DT 01-JUL-1989 (rel. 11, Created)
 DT 01-JUL-1989 (rel. 11, Last sequence update)
 DT 15-DEC-1998 (rel. 37, Last annotation update)
 DE GLUTATHIONE S-TRANSFERASE GST.7 (EC 2.5.1.18) (GST 1-1) (GST CLASS-
 DE MU).
 GN GSTM1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN 11
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88330838; PubMed=3417659;
 RA Pearson W.R., Reinhardt J., Sisk S.C., Anderson K.S., Adler P.N.;
 RT "Tissue-specific induction of murine glutathione transferase mRNAs by
 RT butylated hydroxyanisole."
 RL J. Biol. Chem. 263:13324-13332(1988).
 RN 12
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90094327; PubMed=2689439;
 RA Townsend A.J., Goldsmith M.E., Pickett C.B., Cowan K.H.;
 RT "Isolation, characterization, and expression in Escherichia coli of
 RT two murine Mu class glutathione S-transferase cDNAs homologous to the
 RT rat subunits 3 (Yb1) and 4 (Yb2)."
 RL J. Biol. Chem. 264:21582-21590(1989).
 RN 13
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93290350; PubMed=8512323;
 RA Reinhardt J., Pearson W.R.;
 RT "The structure of two murine class-mu glutathione transferase genes
 RT coordinately induced by butylated hydroxyanisole."
 RL Arch. Biochem. Biophys. 303:383-393(1993).
 RN 14
 RP PRELIMINARY SEQUENCE OF 1-40.
 RX MEDLINE=83109018; PubMed=6822548;
 RA Pearson W.R., Windle J.J., Morrow J.F., Benson A.M., Talalay P.;
 RT "Increased synthesis of glutathione S-transferases in response to
 RT anticarcinogenic antioxidants. Cloning and measurement of messenger
 RT RNA."
 RL J. Biol. Chem. 258:2052-2062(1983).
 RN 15
 RP PRELIMINARY SEQUENCE OF 1-24.
 RX MEDLINE=86042634; PubMed=3864155;
 RA Mannervik B., Alin P., Guthenberg C., Jonsson H., Tahir M.K.,
 RA Warholm M., Joernvall H.;
 RT "Identification of three classes of cytosolic glutathione transferase
 RT common to several mammalian species: correlation between structural
 RT data and enzymatic properties."
 RL Proc. Natl. Acad. Sci. U.S.A. 82:7202-7206(1985).
 RN 16
 RP CHARACTERIZATION.
 RX STRAIN-CD-1; TISSUE-LIVER;
 RX MEDLINE=96189427; PubMed=8605288;
 RA Mitchell A.E., Morlin D., Lane M.W., Jones A.D.;
 RT "Purification, mass spectrometric characterization, and covalent
 RT modification of murine glutathione S-transferases."
 RL Chem. Res. Toxicol. 8:1054-1062(1995).
 CC -1- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
 CC OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
 CC -1- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GUTATHIONE.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- MASS SPECTROMETRY: MW=25838.4; MW_ERR=2; METHOD-ELECTROSPRAY.

CC -1 SIMILARITY: BELONGS TO THE GST SUPERFAMILY. MU FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: J03952; AAA37747.1; -
DR EMBL: J04632; AAA37705.1; -
DR EMBL: L13448; -; NOT_ANNOTATED_CDS.
DR PIR: A20831; A20831.
DR PIR: A28946; A28946.
DR PIR: A24735; A24735.
DR PIR: A34159; A34159.
DR PIR: S33860; S33860.
DR HSSP: P04905; 6GSY.
DR SWISS-2DPAGE: P10649; MOUSE.
DR MGD: MGI:95860; GSTM1.
DR INTERPRO: IPR0003081; -.
DR PRAM: PF00043; GST; 1.
DR PRINTS: PRO1267; GSTTRANSFERASE.
DR TRANSFERASE: Multigene family.
FT INIT_MET 0
SQ SEQUENCE 217 AA; 25839 MW; A1EE3938F590829 CRC64;

Query Match 26.0%; Score 481.5; DB 1; Length 217;
Best Local Similarity 44.4%; Pred. No. 9.6e-33;
Matches 91; Conservative 38; Mismatches 71; Indels 5; Gaps 1;

QY 4 ILGVKKGIVQPTLLILEYLEKEEHEHYERDEG-----DKWRKKKFLGEGFENLPY 58
DB 3 ILGVNNGELHPIMRLLEYDSSYDEKRYTGADPDFDRSQMLEKFLGLDFENLPYL 62
QY 59 IDGVKRLQSMATIRYIADKHNMLGCGPKERAEISMLEGAVIDIRYGVSRISYKDFETL 118
DB 63 IDGSKRTIOSNALIKYLRKHLHLDGETEERIRADIVENQVMDTLMCLCYNDEFEQ 122
QY 119 KVDLSKIPKEMKMFEDRLCKHTYINGDVHPDFMLYDALDVLVYMDPMLDAEPKILVC 178
DB 123 KPEFLKTIPEKMKIYSEFLGKRPWAGDKVTVYDFLAVDILQYRMFEPKCLDAPENLRD 182
QY 179 FKKRIEALPQIDKYLKSSKYTAMPL 203
DB 183 FLAREGLKRIKISAYMKSSRYATPT 207

RESULT 9
GTM1_RAT STANDARD; PRT; 217 AA.
AC P04905;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GLUTATHIONE S-TRANSFERASE YB1 (EC 2.5.1.18) (CHAIN 3) (GST M1-1)
DE (GST CLASS-M0).
GN GSTM1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A. (CLONE PGTR200).
RA MEDLINE=8631282; PubMed=2875437;
RA Lai H.-C., Grove G., Tu C.-P.D.;
RT "Cloning and sequence analysis of a cDNA for a rat liver glutathione
RT S-transferase Yb subunit";
RL Nucleic Acids Res. 14:6101-6114(1986).
RN [2]
RP SEQUENCE FROM N.A. (CLONE PGTA/C44).

RX MEDLINE=86033768; PubMed=3840477;
RA Ding G.-J., Lu A.Y.H., Pickett C.B.;
RT Rat liver glutathione S-transferases. Nucleotide sequence analysis
RT of a Yb1 cDNA clone and prediction of the complete amino acid
RT sequence of the Yb1 subunit";
RL J. Biol. Chem. 260:13268-13271(1985).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=86224097; PubMed=3011803;
RA Ding G.-J., Ding V.D.-H., Rodkey J.A., Bennett C.D., Lu A.Y.H.,
RA Pickett C.B.;
RT "Rat liver glutathione S-transferases. DNA sequence analysis of a Yb2
RT cDNA clone and regulation of the Yb1 and Yb2 mRNAs by
RT phenobarbital";
RL J. Biol. Chem. 261:7952-7957(1986).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=87308179; PubMed=3040722;
RA Chang C., Saltzman A.G., Sorensen N.S., Hlupakka R.A., Liao S.;
RT "Identification of glutathione S-transferase Yb1 mRNA as the
RT androgen-repressed mRNA by cDNA cloning and sequence analysis.";
RL J. Biol. Chem. 262:11901-11903(1987).
RN [5]
RP SEQUENCE OF 1-23.
RC STRAIN-WISTAR; TISSUE-OLFACTORY EPITHELIUM;
RX MEDLINE=93277499; PubMed=8503873;
RA Ben-Arie N., Khen M., Lancel D.;
RT "Glutathione S-transferases in rat olfactory epithelium:
RT purification, molecular properties and odorant biotransformation";
RL Biochem. J. 292:379-384(1993).
RN [6]
RP MUTAGENESIS OF CYS-86.
RX MEDLINE=91354218; PubMed=1883338;
RA Hsieh J.-C., Huang S.-C., Chen W.-L., Lai Y.-C., Tam M.F.;
RT "Cysteine-86 is not needed for the enzymic activity of glutathione S-
RT transferase 3-3";
RL Biochem. J. 278:293-297(1991).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
RX MEDLINE=93041702; PubMed=1420139;
RA Ji X., Zhang P., Armstrong R.N., Gilliland G.L.;
RT "The three-dimensional structure of a glutathione S-transferase from
RT the mu gene class. Structural analysis of the binary complex of
RT isoenzyme 3-3 and glutathione at 2.2-A resolution.";
RL Biochemistry 31:10169-10184(1992).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RA Fu J.-H., Rose J., Tam M.F., Wang B.-C.;
RT "New crystal forms of a mu-class glutathione S-transferase from rat
RT liver";
RL Acta Crystallogr. D 50:219-224(1994).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=94153886; PubMed=8110735;
RA Ji X., Johnson W.W., Sesay M.A., Dickert L., Prasad S.M., Ammon H.L.,
RA Armstrong R.N., Gilliland G.L.;
RT "Structure and function of the xenobiotic substrate binding site of a
RT glutathione S-transferase as revealed by x-ray crystallographic
RT analysis of product complexes with the diastereomers of 9-(S-
RT glutathionyl)-10-hydroxy-9,10-dihydrophenanthrene.";
RL Biochemistry 33:1043-1052(1994).
CC -1- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
CC OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
CC -1- FUNCTION: THE OLFACTORY GST MAY BE CRUCIAL FOR THE ACTIVITY OF THE
CC OLFACTORY PROCESS.
CC -1- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GLUTATHIONE.
CC -1- SUBUNIT: HOMODIMER OR HETERODIMER.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- MISCELLANEOUS: RAT LIVER GLUTATHIONE S-TRANSFERASES (GST) CONSIST
CC OF BINARY COMBINATIONS OF THREE MAJOR CLASSES OF SUBUNITS
CC DESIGNATED: YA (25 KDA), YB (27 KDA) AND YC (28 KDA).
CC -1- MISCELLANEOUS: YB SUBCLASS SELECTIVELY BINDS STEROID HORMONES.
CC -1- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. MU FAMILY.

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
CC
CC EMBL; X04229; CAA27811.1; -
CC DR EMBL; M11719; AAA41287.1; -
CC DR EMBL; J02810; AAA41293.1; -
CC DR PIR; A24085; A24085.
CC DR PIR; A25510; A25510.
CC DR PIR; A29794; A29794.
CC DR PIR; S17167; S17167.
CC DR PDB; 1G5B; 31-OCT-93.
CC DR PDB; 1G5C; 31-OCT-93.
CC DR PDB; 2GST; 31-OCT-93.
CC DR PDB; 3GST; 31-JAN-94.
CC DR PDB; 4GST; 31-OCT-93.
CC DR PDB; 5GST; 31-OCT-93.
CC DR PDB; 6GST; 08-NOV-96.
CC DR PDB; 6GSU; 08-NOV-96.
CC DR PDB; 6GSV; 08-NOV-96.
CC DR PDB; 6GSW; 08-NOV-96.
CC DR PDB; 6GSX; 08-NOV-96.
CC DR PDB; 6GSY; 08-NOV-96.
CC DR PDB; 5FWG; 27-JAN-99.
CC DR INTERPRO; IPR000521; -
CC DR INTERPRO; IPR003081; -
CC DR PFAM; PF00043; GST; 1.
CC DR PRINTS; PRO1267; GSTRNSFRASEM.
CC KW Transferrase; Multigene family; 3D-structure; Olfaction.
CC FT INIT MET 0 0
CC FT MUTAGEN 86 86 C->S; NO CHANGE IN ACTIVITY.
CC FT CONFLICT 168 168 I-> N (IN REF. 3).
CC FT CONFLICT 198 199 KS -> NC (IN REF. 2).
CC FT STRAND 2 7
CC FT TURN 11 13
CC FT HELIX 14 22
CC FT TURN 23 24
CC FT STRAND 27 32
CC FT TURN 37 39
CC FT HELIX 43 46
CC FT TURN 47 50
CC FT STRAND 61 64
CC FT TURN 65 66
CC FT STRAND 67 70
CC FT HELIX 72 82
CC FT TURN 83 84
CC FT HELIX 90 114
CC FT TURN 115 115
CC FT TURN 117 118
CC FT HELIX 119 128
CC FT TURN 129 129
CC FT HELIX 130 141
CC FT TURN 142 143
CC FT TURN 154 154
CC FT HELIX 155 169
CC FT TURN 171 176
CC FT HELIX 178 188
CC FT TURN 189 189
CC FT HELIX 191 196
CC FT TURN 197 198
CC FT TURN 200 201
CC FT TURN 210 211
CC SEQUENCE 217 AA; 25782 MW; 2ACE8D49DA785118 CRC64;
CC
Query Match 26.0%; Score 481.5; DB 1; Length 217;
Best Local Similarity 44.4%; Pred. No. 9.6e-33;
Matches 91; Conservative 38; Mismatches 71; Indels 5; Gaps 14;

```

```

OY 4 ILGVKRIKGLVQPTLLLEYLEEKKEEHLNDEDEG-----DKWRNKKPEELGLEFPNLPY 58
OS |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3 ILGVNNGVGLTHPIPLLELTETDSSTEEEKRYAMGADPDYDRSOWLMEKRTLGDLPFNLPYL 62
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 59 IDGVKLTQSMANITIVYADKHNMLGCGCKPERAETSMLEGCVALDIRVGYSRVIAYSKDFEPL 118
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 63 IDGSKRIITQSMALIMRYLARKHHLGCETEDEERIRADIVENQWVDNMNMOILMICYNDFEKO 122
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 119 KVDPLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDMYLTDALDVLVMDPMLCDAEPKVLVC 178
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 123 KPEEFKLTPEKKMKLTSEELGKRPMPAGKRVYVDFLAVDILDQVHIFEPKCDLAEFNLKD 182
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 179 FKRIEALPQIDKYLKSSKYIAMP 203
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 183 FLAREEGKIKISAVYKSSRYLSTPI 207
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 10
GTM4_HUMAN
ID GTM4_HUMAN STANDARD: PRT: 218 AA.
AC 003013.1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE GLUTATHIONE S-TRANSFERASE MU 4 (BC 2.5.1.18) (GSTM4-4) (GTS-MU2)
DE (GST CLASS-MU).
GN GSTM4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
RN [1]
RP SEQUENCE FROM N.A..
RX MEDLINE=93352467; Pubmed-8349586;
RA Comstock K.E., Johnson K.J., Rittenberg D., Henner W.D.;
RT "Isolation and analysis of the gene and cDNA for a human Mu class
RT glutathione S-transferase, GSTM4.";
RL J. Biol. Chem. 268:16958-16965(1993).
RN [2]
RP SEQUENCE FROM N.A..
RX MEDLINE=93228631; Pubmed-8471052;
RA Zhong S., Spurr N.K., Hayes J.D., Wolf C.R.;
RT "Deduced amino acid sequence, gene structure and chromosomal location
RT of a novel human class Mu glutathione S-transferase, GSTM4.";
RL Biochem. J. 291:41-50(1993).
RN [3]
RP SEQUENCE OF 39-160 FROM N.A..
RX TISSUE=Lymphocytes;
RC MEDLINE=91174774; Pubmed=2006920;
RA Taylor J.B., Oliver J., Sherrington R., Pemble S.E.;
RT "Structure of human glutathione S-transferase class Mu genes.";
RL Biochem. J. 274:587-593(1991).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=94263230; Pubmed-8203914;
RA Comstock K.E., Widersten M., Hao X.Y., Henner W.D., Manervik B.;
RT "A comparison of the enzymatic and physicochemical properties of
RT human glutathione transferase M4-4 and three other human Mu class
RT enzymes.";
RL Arch. Biochem. Biophys. 311:487-495(1994).
CC -1- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
CC OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
CC ACTIVE ON 1-CHLORO-2,4-DINITROBENZENE.
CC -1- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GUTATHIONE.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN A WIDE VARIETY OF TISSUES.
CC -1- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. MU FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```



```
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M63509; AAA60963.1; -.
DR PIR: A39375; A39375.
DR PDB: 1HNA; 31-JAN-94.
DR PDB: 1HNB; 31-JAN-94.
DR PDB: 1HNC; 31-JAN-94.
DR PDB: 2GTU; 02-MAR-99.
DR MIM: 138380; -.
DR INTERPRO: IPR000521; -.
DR PFAM: PF00043; GST; 1.
DR PRINTS: PRO1267; GSTRNSFRASEM.
DR TRANSFERASE; Multigene family; 3D-structure.
KW INIT_MET 0
FT STRAND 2 5
FT TURN 7 7
FT TURN 11 12
FT TURN 13 22
FT TURN 23 24
FT STRAND 27 29
FT STRAND 32 32
FT STRAND 35 35
FT STRAND 38 39
FT TURN 41 41
FT TURN 43 49
FT TURN 50 52
FT STRAND 61 64
FT TURN 65 66
FT STRAND 67 70
FT TURN 72 81
FT TURN 82 84
FT TURN 90 113
FT TURN 114 115
FT TURN 117 118
FT TURN 119 127
FT TURN 128 129
FT TURN 130 141
FT TURN 142 143
FT STRAND 146 146
FT TURN 147 147
FT STRAND 148 149
FT TURN 150 150
FT STRAND 151 151
FT TURN 154 169
FT TURN 171 176
FT TURN 178 189
FT TURN 191 197
FT TURN 198 198
FT TURN 214 215
SO SEQUENCE 217 AA; 25613 MW; 25603A909482CA39 CRC64;
```

Query Match 25.6%; Score 473.5; DB 1; Length 217;
Best Local Similarity 44.3%; Pred. No. 4.4e-32;
Matches 94; Conservative 35; Mismatches 78; Indels 5; Gaps 1;

```
OY 5 LGYWKIKGLVQPTRLLEYLEEKEEHLVERDEG-----DKWRNKKFELGLEFPNLPYYI 59
DB 4 LGYWNIRGLAHSTRILEETDSSYEKKRYTMGAPYDRSQWLNKFKLGLDFPNLPYLI 63
OY 60 DGDVKTLOSMAITRYTADKHNMLGCPKERAETSMLEGAVALDIRYGVSRITAYSKDEFETLK 119
DB 64 DGHKTKTOSNAITRYTARKHNMLGSESEKEQIRIEDILENQEMDSRMQLAKCTPPDEFK 123
OY 120 VDFLSKLPKMLKMFEDRLCHKTYLNGDHYTHPDMYLDALDVLVLYMDPMCLDAFPKLYCF 179
```

```
DB 124 PERIALPPEMLKLYSOFLOKOPFLCDKITFEVFIAYDVLEARNQVEPSCDLAFPIKLDF 183
OY 180 KRRIEAIPOIDKLYSKSYIAMPLQGMATFG 211
DB 184 ISRFEGLEKISAYMKSSRFLPRVPTKMAVG 215

RESULT 13
GTMU_CAVPO
ID GTMU_CAVPO STANDARD; PRT; 217 AA.
AC P16413;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE GLUTATHIONE-S-TRANSFERASE B (EC 2.5.1.18) (GST B) (GST CLASS-MU).
GN GSTM1.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriognathi; Cavidae; Cavia.
RN [1]
RP SEQUENCE.
RC TISSUE=LIVER;
RX MEDLINE=90236961; PubMed=2332413;
RA Kamei K., Oshino R., Hara S.;
RT "Amino acid sequence of glutathione S-transferase b from guinea pig
  liver."
RL J. Biochem. 107:111-117(1990).
CC -1- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
CC OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
CC -1- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GUTATHIONE.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. MU FAMILY.
DR PIR: JX0095; JX0095.
DR HSSP: P04905; 6GSY.
DR INTERPRO: IPR000521; -.
DR INTERPRO: IPR003081; -.
DR PFAM: PF00043; GST; 1.
DR PRINTS: PRO1267; GSTRNSFRASEM.
KW TRANSFERASE; Multigene family.
SQ SEQUENCE 217 AA; 25719 MW; D29F7951D4E9365E CRC64;
```

Query Match 25.6%; Score 473.5; DB 1; Length 217;
Best Local Similarity 45.1%; Pred. No. 4.4e-32;
Matches 92; Conservative 34; Mismatches 73; Indels 5; Gaps 1;

```
OY 5 LGYWKIKGLVQPTRLLEYLEEKEEHLVERDEG-----DKWRNKKFELGLEFPNLPYYI 59
DB 4 LGYWNIRGLHPIRLILEYNSGYEKKRYNMGAPYDRSQWLNKFKLGLDFPNLPYLI 63
OY 60 DGDVKTLOSMAITRYTADKHNMLGCPKERAETSMLEGAVALDIRYGVSRITAYSKDEFETLK 119
DB 64 DGHKTKTOSNAITRYTARKHNMLCGVTEETIRMDILENQYMDIRMDIMLCYSPDEFQOK 123
OY 120 VDFLSKLPKMLKMFEDRLCHKTYLNGDHYTHPDMYLDALDVLVLYMDPMCLDAFPKLYCF 179
DB 124 AEFLEGIPTDKMLFSOFLGLPMPFAGKLTLYVDFLAVLDVYRMLEPMLKDAFPNKLDF 183
OY 180 KRRIEAIPOIDKLYSKSYIAMPL 203
DB 184 ISRFEGLEKISAYMKSSRFLPRPL 207
```

RESULT 14
GTMU_CRILLO
ID GTMU_CRILLO STANDARD; PRT; 217 AA.
AC 000285;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GLUTATHIONE S-TRANSFERASE Y1 (EC 2.5.1.18) (CHAIN 3) (GST CLASS-MU).

```

OS Cricetus longicaudatus (Long-tailed hamster) (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
RN [1]
RP SEQUENCE FROM N.A. PubMed=2226437;
RX MEDLINE=91031445; Hyrien O., Debatisse M., Buttin G.;
RA de Saint Vincent B.R., "Complication of mu class glutathione S-transferase genes and an
RT adenylyate deaminase gene in coformycin-resistant Chinese hamster
RT fibroblasts."
RL Eur. J. Biochem. 193;19-24(1990).
CC -1- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
CC OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
CC -1- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GUTATHIONE.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. MU FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X57489; CAA40726.1; -.
DR PIR; S13202; S13202.
DR HSSP; P04905; 6GST.
DR INTERPRO: IPR000521; -.
DR INTERPRO: IPR003081; -.
DR PFAM; PF00043; GST; 1.
DR PRINTS; PR01267; GSTRNSFRASEM.
DR Transferase; Multigene family.
KW INIT MET 0
FT SEQUENCE 217 AA; 25688 MW; 34D8A8FBDD8627C0 CRC64;
SO

Query Match 25.3%; Score 468.5; DB 1; Length 217;
Best Local Similarity 43.4%; Pred. No. 1.1e-31;
Matches 89; Conservative 36; Mismatches 75; Indels 5; Gaps 1;

QY 4 ILGVWIKIGVOPTRLLLEYLEEYEHLYERDEG-----DKRNKKFELGLEPMLPYI 58
DB 3 ILGIVWYKGLTNPRIKLLLEITDSYEKKYTMGDAPDSRQWLNKFKLGDFPMLPYL 62
QY 59 IDGDVKLTOSMAITRYIADKHNMLGGCPKRAEISMEGAVIDIRYGVSRIVASKPEFL 118
DB 63 IDGSHKITQSNALIRYIARKHNLGCTEEERIVYDIVENQAMDTROMLIMLCYNPDEKQ 122
QY 119 KVDFLSLPEMLKMFEDRCLCHKTYLNGDHYTHDPMLYDALDVVLVMDPCLDAFRLKVC 178
DB 123 KPEFLKTIPIKRMWSEFLGRFWRPADKVTLLCGFLAYDVLDOYQFEPKCLDPFRLND 182
QY 179 EKRIEAIPOIDKYLYKSSKYIAMPL 203
DB 183 FLAFEGKLKRTISAVMTKSRFLRRI 207

```

```

RP SEQUENCE FROM N.A.
RC TISSUE=SMOOTH MUSCLE;
RX MEDLINE=92335246; PubMed=1631097;
RA Fan W.M., Trifiletti R., Norris J.S., Cooper T.M.;
RT "Cloning of a mu-class glutathione S-transferase gene and
RT identification of the glucocorticoid regulatory domains in its 5'
RT flanking sequence."
RL Proc. Natl. Acad. Sci. U.S.A. 89;6104-6108(1992).
RN [2]
RN SEQUENCE FROM N.A.
RP TISSUE=SMOOTH MUSCLE;
RX MEDLINE=92049380; PubMed=1944302;
RA Norris J.S., Schwartz D.A., Macleod S.L., Fan W.M., O'Brien T.J.,
RA Harris S.E., Trifiletti R., Cornett L.E., Cooper T.M., Levi W.M.,
RA Smith R.G.;
RT "Cloning of a mu-class glutathione S-transferase complementary DNA
RT and characterization of its glucocorticoid inducibility in a smooth
RT muscle tumor cell line."
RL Mol. Endocrinol. 5;979-986(1991).
CC -1- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
CC OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
CC -1- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GUTATHIONE.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. MU FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M59772; AAA37075.1; -.
DR EMBL; X61033; CAA43368.1; -.
DR PIR; A23732; A23732.
DR PIR; S21908; S21908.
DR HSSP; P04905; 6GST.
DR INTERPRO: IPR000521; -.
DR INTERPRO: IPR003081; -.
DR PFAM; PF00043; GST; 1.
DR PRINTS; PR01267; GSTRNSFRASEM.
DR Transferase; Multigene family.
KW INIT MET 0
FT SEQUENCE 217 AA; 25558 MW; C59812AA430B6375 CRC64;
SO

Query Match 25.3%; Score 468.5; DB 1; Length 217;
Best Local Similarity 44.7%; Pred. No. 1.1e-31;
Matches 89; Conservative 37; Mismatches 68; Indels 5; Gaps 1;

QY 5 LGVWIKIGVOPTRLLLEYLEEYEHLYERDEG-----DKRNKKFELGLEPMLPYI 59
DB 4 LGVWIRGLAHRIKLLLEITDSYEKKYTMGDAPNFRSQWLNKFKLGDFPMLPYL 63
QY 60 DGDVKLTOSMAITRYIADKHNMLGGCPKRAEISMEGAVIDIRYGVSRIVASKPEFLK 119
DB 64 DGSHKITQSNALIRYIARKHNLGCTEEERIQDLTDENQAMDTROMLIMVCCSPDEKQ 123
QY 120 VDFLSLPEMLKMFEDRCLCHKTYLNGDHYTHDPMLYDALDVVLVMDPCLDAFRLKVC 179
DB 124 PEYLEDLPKRMILYSEFLGRFWRPADKVTYDFLLIYDVLQDHRIFAFKCLDAFRLKDF 183
QY 180 KRIEAIPOIDKYLYKSSKY 198
DB 184 LAFEGKLKRTISAVMTKSRF 202

```

Search completed: March 18, 2001, 05:38:37
Job time: 480 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2001, 05:29:32 ; Search time 74.32 Seconds
(without alignments)
540.935 Million cell updates/sec

Title: US-09-402-488a-2
Perfect score: 1650
Sequence: 1 MSPILGWKIKGLVQPTRL.....PKQSHNDGFEELPEEXLQ 343

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues
Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1244.5	67.3	245	2	009193
2	1226	66.3	262	2	009195
3	1226	66.3	268	2	009194
4	1202.5	65.0	243	2	009192
5	1202.5	65.0	244	2	009191
6	1049	56.7	218	5	094745
7	718	38.8	218	5	025595
8	698	37.7	218	5	09XYL9
9	512.5	27.2	223	5	09N0V4
10	502.5	25.9	218	11	035660
11	478.5	25.8	218	6	09TSM5
12	477.5	25.6	218	6	09TSM4
13	474.5	25.6	219	5	027653
14	473.5	25.6	219	5	016058
15	464.5	25.1	218	11	09WU21
16	448.5	24.2	225	4	060550
17	436.5	23.6	195	4	005465
18	434.5	23.5	125	11	09Z1B2

20	407.5	22.0	188	6	09WZB4	09WZB4 capra hircu
21	377.5	20.4	219	5	090582	090582 psoroptes o
22	288.5	15.6	125	6	029583	029583 sus scrofa
23	256.5	13.9	208	5	002636	002636 brugia mala
24	247.5	13.4	208	5	027711	027711 onchocerca
25	247	13.4	84	5	007557	007557 hirudinaria
26	246	13.3	210	13	P81942	P81942 bufo bufo
27	240	13.0	210	4	Q15690	Q15690 homo sapien
28	240	13.0	210	4	000460	000460 homo sapien
29	236	12.8	180	5	09NHB2	09NHB2 mytilus edu
30	233	12.6	209	6	09TTY8	09TTY8 capra hircu
31	230	12.4	84	5	007558	007558 hirudinaria
32	213.5	11.5	216	13	Q9W647	Q9W647 oncorhynch
33	212.5	11.5	82	4	Q9UE37	Q9UE37 homo sapien
34	203	11.0	366	6	09N1P5	09N1P5 bubalus bub
35	197.5	10.7	345	6	027951	027951 bos taurus
36	192.5	10.4	208	5	P91505	P91505 caenorhabd
37	191.5	10.4	226	11	Q90LX3	Q90LX3 rattus norv
38	189.5	10.2	221	11	P70686	P70686 mesocricetu
39	187	10.1	210	5	Q9N4X8	Q9N4X8 caenorhabd
40	186.5	10.1	222	6	09N2J6	09N2J6 ovis aries
41	182.5	9.9	208	5	061750	061750 caenorhabd
42	182.5	9.9	223	6	Q29057	Q29057 sus scrofa
43	180.5	9.8	199	11	Q90HE7	Q90HE7 mus musculu
44	179.5	9.7	222	4	Q9NTY6	Q9NTY6 homo sapien
45	179.5	9.7	222	6	Q9XS30	Q9XS30 ovis aries

ALIGNMENTS

```
RESULT 1
ID 009193 PRELIMINARY; PRT; 245 AA.
AC 009193;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE GLUTATHIONE S-TRANSFERASE.
GN GST.
OS Escherichia coli.
OG Plasmid pGEX-6P-3.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN 111
RP SEQUENCE FROM N.A.
RX MEDLINE=88329742; PubMed=3047011;
RA Smith D.B., Johnson K.S.;
RT "Single-step purification of polypeptides expressed in Escherichia
RT coli as fusions with glutathione S-transferase.";
RL Gene 67:31-40(1988).
RN 121
RP SEQUENCE OF 221-227 FROM N.A.
RX MEDLINE=90264387; PubMed=2160953;
RA Cordingley M.G., Callahan P.L., Sardana V.V., Garsky V.M.,
RA Colombo R.J.;
RT "Substrate requirements of human rhinovirus 3C protease for peptide
RT cleavage in vitro.";
RL J. Biol. Chem. 265:9062-9065(1990).
RN 131
RP SEQUENCE FROM N.A.
RA Dunst R.W., Bell P.A., English J.D.;
RA Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U78874; AAB37352.1; -.
DR HSSP; P08515; IGNE.
DR INTERPRO; IPR000521; -.
DR PFM; PF00043; GST_1.
KW Transferase; Plasmid.
SQ SEQUENCE 245 AA; 28327 MW; D51470B8A48C6CC2 CRC64;
```

Query Match

67.3%; Score 1244.5; DB 2; Length 245;

Best Local Similarity 97.1%; Pred. No. 2.1e-94;
Matches 234; Conservative 0; Mismatches 2; Indels 5; Gaps 1;

```
OY 1 MSPILGYMKIKGLVPTLLLEYLEEKYEENLYERDEGDKMNKKFELGLEFPNLPYYID 60
DB 1 MSPILGYMKIKGLVPTLLLEYLEEKYEENLYERDEGDKMNKKFELGLEFPNLPYYID 60
OY 61 GVKLTQSMATIRYIADKHNMLGGCPKERAETSMLEGAVALDIRYGSRIAYSKDEFTLV 120
DB 61 GVKLTQSMATIRYIADKHNMLGGCPKERAETSMLEGAVALDIRYGSRIAYSKDEFTLV 120
OY 121 DFLSKLPEMLKMFEDRLCHKTYLNGDHYTHPDPMLYDALDVLVYMDPMCLDAFPKLVCFK 180
DB 121 DFLSKLPEMLKMFEDRLCHKTYLNGDHYTHPDPMLYDALDVLVYMDPMCLDAFPKLVCFK 180
OY 121 DFLSKLPEMLKMFEDRLCHKTYLNGDHYTHPDPMLYDALDVLVYMDPMCLDAFPKLVCFK 180
DB 121 DFLSKLPEMLKMFEDRLCHKTYLNGDHYTHPDPMLYDALDVLVYMDPMCLDAFPKLVCFK 180
OY 181 KRIEAIPOIDKYLKSSKYIAMPLOGMOATFGGSDHPKPSDLPVPRGSP 235
DB 181 KRIEAIPOIDKYLKSSKYIAMPLOGMOATFGGSDHPKPSDLPVPRGSP 240
OY 236 R 236
DB 241 R 241
```

RESULT 2
009195 PRELIMINARY: PRT: 262 AA.
AC 009195:
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE GST-6HIS.
GN GST.
OS Escherichia coli.
OG Plasmid pGEX-2T-6His-PL2.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RA Hipskind R.H., Delaney P.;
RL Submitted (Feb-1997) to the EMBL/GenBank/DBJ databases.
RN [12]
RP SEQUENCE FROM N.A.
RA Canill M.A.;
RL Submitted (Jan-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U84572; AAB41883.1; -
DR HSSP: P08515; IGNE.
DR INTERPRO: IPR000521; -
DR PFM: PF00043; GST; 1.
SQ SEQUENCE 262 AA; 30392 MW; BA4C4593CB20226E CRC64;

Query Match 66.3%; Score 1226; DB 2; Length 262;
Best Local Similarity 100.0%; Pred. No. 7.6e-93;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 MSPILGYMKIKGLVPTLLLEYLEEKYEENLYERDEGDKMNKKFELGLEFPNLPYYID 60
DB 1 MSPILGYMKIKGLVPTLLLEYLEEKYEENLYERDEGDKMNKKFELGLEFPNLPYYID 60
OY 61 GVKLTQSMATIRYIADKHNMLGGCPKERAETSMLEGAVALDIRYGSRIAYSKDEFTLV 120
DB 61 GVKLTQSMATIRYIADKHNMLGGCPKERAETSMLEGAVALDIRYGSRIAYSKDEFTLV 120
OY 121 DFLSKLPEMLKMFEDRLCHKTYLNGDHYTHPDPMLYDALDVLVYMDPMCLDAFPKLVCFK 180
DB 121 DFLSKLPEMLKMFEDRLCHKTYLNGDHYTHPDPMLYDALDVLVYMDPMCLDAFPKLVCFK 180
OY 121 DFLSKLPEMLKMFEDRLCHKTYLNGDHYTHPDPMLYDALDVLVYMDPMCLDAFPKLVCFK 180
DB 121 DFLSKLPEMLKMFEDRLCHKTYLNGDHYTHPDPMLYDALDVLVYMDPMCLDAFPKLVCFK 180
OY 181 KRIEAIPOIDKYLKSSKYIAMPLOGMOATFGGSDHPKPSDLPVPRGSP 227
DB 181 KRIEAIPOIDKYLKSSKYIAMPLOGMOATFGGSDHPKPSDLPVPRGSP 227
```

```
RESULT 3  
009194 PRELIMINARY: PRT: 268 AA.  
AC 009194:  
DT 01-JUL-1997 (TREMBLrel. 04, Created)  
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
DE GST-6HIS.  
GN GST.  
OS Escherichia coli.  
OG Plasmid pGEX-2T-6H.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Canill M.A.;  
RL Submitted (Jan-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL: U84571; AAB41882.1; -  
DR HSSP: P08515; IGNE.  
DR INTERPRO: IPR000521; -  
DR PFM: PF00043; GST; 1.  
SQ SEQUENCE 268 AA; 31000 MW; 85820375FE60E625 CRC64;
```

Query Match 66.3%; Score 1226; DB 2; Length 268;
Best Local Similarity 100.0%; Pred. No. 7.8e-93;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 MSPILGYMKIKGLVPTLLLEYLEEKYEENLYERDEGDKMNKKFELGLEFPNLPYYID 60
DB 1 MSPILGYMKIKGLVPTLLLEYLEEKYEENLYERDEGDKMNKKFELGLEFPNLPYYID 60
OY 61 GVKLTQSMATIRYIADKHNMLGGCPKERAETSMLEGAVALDIRYGSRIAYSKDEFTLV 120
DB 61 GVKLTQSMATIRYIADKHNMLGGCPKERAETSMLEGAVALDIRYGSRIAYSKDEFTLV 120
OY 121 DFLSKLPEMLKMFEDRLCHKTYLNGDHYTHPDPMLYDALDVLVYMDPMCLDAFPKLVCFK 180
DB 121 DFLSKLPEMLKMFEDRLCHKTYLNGDHYTHPDPMLYDALDVLVYMDPMCLDAFPKLVCFK 180
OY 121 DFLSKLPEMLKMFEDRLCHKTYLNGDHYTHPDPMLYDALDVLVYMDPMCLDAFPKLVCFK 180
DB 121 DFLSKLPEMLKMFEDRLCHKTYLNGDHYTHPDPMLYDALDVLVYMDPMCLDAFPKLVCFK 180
OY 181 KRIEAIPOIDKYLKSSKYIAMPLOGMOATFGGSDHPKPSDLPVPRGSP 227
DB 181 KRIEAIPOIDKYLKSSKYIAMPLOGMOATFGGSDHPKPSDLPVPRGSP 227
```

RESULT 4
009192 PRELIMINARY: PRT: 243 AA.
AC 009192:
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE GLUTATHIONE S-TRANSFERASE.
GN GST.
OS Escherichia coli.
OG Plasmid pGEX-6P-2.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=88329742; Pubmed=3047011;
RX Smith D.B., Johnson K.S.;
RT "Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase.";
RL Gene 67:31-40(1988).
RN [12]
RP SEQUENCE OF 221-227 FROM N.A.
RX MEDLINE=90264387; Pubmed=2160953;
RA Cordingley M.G., Callahan P.L., Sardana V.V., Garsky V.M.,
RA Colonna R.J.;

RT "Substrate requirements of human rhinovirus 3C protease for peptide cleavage in vitro.";
 RL J. Biol. Chem. 265:9062-9065(1990).
 RP [3]
 RN SEQUENCE FROM N.A.
 RA Dunst R.W., Bell P.A., English J.D.;
 RL EMBL: 78873; AAB37349.1; -
 DR HSSP: P08515; IGNE.
 DR INTERPRO: IPR000521; -
 DR PFAM: PF00043; GST; 1.
 DR Transferase; Plasmid
 KW SEQUENCE 243 AA; 27908 MW; 0003C434CD76C3A6 CRC64;
 SQ
 Query Match 65.0%; Score 1202.5; DB 2; Length 243;
 Best Local Similarity 94.2%; Pred. No. 5.8e-91;
 Matches 228; Conservative 1; Mismatches 6; Indels 7; Gaps 2;
 QY 1 MSPILGYWKIKGLVQPTRLLEYLEEKYEHLIERDEGDKWRKKFELGLEFPNLPYYID 60
 DB 1 MSPILGYWKIKGLVQPTRLLEYLEEKYEHLIERDEGDKWRKKFELGLEFPNLPYYID 60
 QY 61 GGVKLTQSMALIRYIADKHNMLGGCCKERAETSMLEGAVIDIRYGSRTAYSDFTLKY 120
 DB 61 GGVKLTQSMALIRYIADKHNMLGGCCKERAETSMLEGAVIDIRYGSRTAYSDFTLKY 120
 QY 121 DFLSKPEMLKMFEDRLCHKTYLNGDVTHTPDMLYDALDVLYMPMCLDAFPKLVCFK 180
 DB 121 DFLSKPEMLKMFEDRLCHKTYLNGDVTHTPDMLYDALDVLYMPMCLDAFPKLVCFK 180
 QY 181 KRIEAIPOIDKYLKSSKYIAMPLOGWQATFGGDDHPKSDLV-----PRGSPNSRVDSG 235
 DB 181 KRIEAIPOIDKYLKSSKYIAMPLOGWQATFGGDDHPKSDLV-----IPGST 238
 QY 236 RA 237
 DB 239 RA 240
 RESULT 5
 CO9191 PRELIMINARY; PRT; 244 AA.
 ID 009191;
 AC 009191;
 DT 01-JUL-1997 (TREMBLrel. 04, Created)
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE GLUTATHIONE S-TRANSFERASE.
 GN GST.
 OS Escherichia coli.
 OG Plasmid pGEX-6P-1.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=88329742; PubMed=3047011;
 RA Smith D.B., Johnson K.S.;
 RT "Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase.";
 RL Gene 67:31-40(1988).
 RL [2]
 RN SEQUENCE OF 221-227 FROM N.A.
 RX MEDLINE=90264387; PubMed=2160953;
 RA Cordingley M.G., Callahan P.L., Sardana V.V., Garsky V.M.,
 RA Colomo R.J.;
 RT "Substrate requirements of human rhinovirus 3C protease for peptide cleavage in vitro.";
 RL J. Biol. Chem. 265:9062-9065(1990).
 RN [3]
 RN SEQUENCE FROM N.A.
 RA Dunst R.W., Bell P.A., English J.D.;
 RL Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.

DR EMBL: 78872; AAB37346.1; -
 DR HSSP: P08515; IGNE.
 DR INTERPRO: IPR000521; -
 DR PFAM: PF00043; GST; 1.
 DR Transferase; Plasmid
 KW SEQUENCE 244 AA; 28430 MW; 7461B31F17027042 CRC64;
 SQ
 Query Match 65.0%; Score 1202.5; DB 2; Length 244;
 Best Local Similarity 93.8%; Pred. No. 5.8e-91;
 Matches 228; Conservative 1; Mismatches 5; Indels 9; Gaps 2;
 QY 1 MSPILGYWKIKGLVQPTRLLEYLEEKYEHLIERDEGDKWRKKFELGLEFPNLPYYID 60
 DB 1 MSPILGYWKIKGLVQPTRLLEYLEEKYEHLIERDEGDKWRKKFELGLEFPNLPYYID 60
 QY 61 GGVKLTQSMALIRYIADKHNMLGGCCKERAETSMLEGAVIDIRYGSRTAYSDFTLKY 120
 DB 61 GGVKLTQSMALIRYIADKHNMLGGCCKERAETSMLEGAVIDIRYGSRTAYSDFTLKY 120
 QY 121 DFLSKPEMLKMFEDRLCHKTYLNGDVTHTPDMLYDALDVLYMPMCLDAFPKLVCFK 180
 DB 121 DFLSKPEMLKMFEDRLCHKTYLNGDVTHTPDMLYDALDVLYMPMCLDAFPKLVCFK 180
 QY 181 KRIEAIPOIDKYLKSSKYIAMPLOGWQATFGGDDHPKSDLV-----PRGSPNSRVDSG 235
 DB 181 KRIEAIPOIDKYLKSSKYIAMPLOGWQATFGGDDHPKSDLV-----EFGP 236
 QY 236 RAE 238
 DB 237 RLE 239
 RESULT 6
 Q94745 PRELIMINARY; PRT; 218 AA.
 ID 094745;
 AC 094745;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE 26 KDA GLUTATHIONE S-TRANSFERASE.
 GN Schistosoma mekongi.
 OS Schistosoma mekongi.
 OC Strigoida; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea;
 OC Eukaryota; Metazoa; Platyhelminthes; Schistosomatidae; Schistosoma.
 OX NCBI_TaxID=38744;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=THAILAND;
 RX MEDLINE=98222395; PubMed=9561610;
 RA Grams S.V., Grams R., Korge G., Viyanant V., Upatham S.;
 RT "Cloning and sequence analysis of the 26 kda glutathiones-transferase gene of Schistosoma mekongi.";
 RL Southeast Asian J. Trop. Med. Public Health 28:570-574(1997).
 DR EMBL: Y07663; CAA68944.1; -
 DR HSSP: P08515; IGTA.
 DR INTERPRO: IPR000521; -
 DR PFAM: PF00043; GST; 1.
 DR Transferase;
 KW SEQUENCE 218 AA; 25486 MW; D3080620B19DE23A CRC64;
 SQ
 Query Match 56.7%; Score 1049; DB 5; Length 218;
 Best Local Similarity 89.4%; Pred. No. 1.9e-78;
 Matches 195; Conservative 11; Mismatches 12; Indels 0; Gaps 0;
 QY 1 MSPILGYWKIKGLVQPTRLLEYLEEKYEHLIERDEGDKWRKKFELGLEFPNLPYYID 60
 DB 1 MSPILGYWKIKGLVQPTRLLEYLEEKYEHLIERDEGDKWRKKFELGLEFPNLPYYID 60
 QY 61 GGVKLTQSMALIRYIADKHNMLGGCCKERAETSMLEGAVIDIRYGSRTAYSDFTLKY 120
 DB 61 GGVKLTQSMALIRYIADKHNMLGGCCKERAETSMLEGAVIDIRYGSRTAYSDFTLKY 120

```

OY 121 DFLSKLPKMFEDRLCKHTYLNCDHVTHPDFMLYDALDVLVYMDPMLDAPFKLVCFK 180
DB 121 DFLNLPKMFEDRLCKHTYLNCDHVTHPDFMLYDALDVLVYMDRCKLDAFPKLVCFK 180
OY 181 KRIEAIPOIDKYLKSSKYIAMPLOGQATFGGDDHPK 218
DB 181 KRIENIPOINEYLKSSKYIEMPLQWQATFGGDDHPK 218

RESULT 7
OY 025595 PRELIMINARY; PRT: 218 AA.
AC 025595;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1999 (Tremblrel. 12, last sequence update)
DT 01-MAY-2000 (Tremblrel. 13, last annotation update)
DE PUTATIVE GLUTATHIONE TRANSFERASE.
OS Clonorchis sinensis.
OC Eukaryota; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea;
OC Opisthorchida; Opisthorchiata; Opisthorchioidea; Opisthorchiidae;
OC Clonorchis.
OX NCBI_TaxID=79923;
RN [1]
RP SEQUENCE FROM N.A.
RA Hong S.-J., Lee D.-H.;
RT "Cloning and over expression of 26 kDa glutathione S-transferase from
RT Clonorchis sinensis."
RL Thesis (1996), Parasitology, Chung-Ang University College of Medicine,
RL Seoul.
RN [2]
RP SEQUENCE FROM N.A.
RA Hong S.-J.;
RT Submitted (APR-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: L47992; AAB46369.3;
DR HSSP: P31670; 1FHE.
DR INTERPRO: IPR000521;
DR PFM: PFO0043; GST; 1.
KM Transferase.
SQ SEQUENCE 218 AA; 25038 MW; 7CB17C7B837A0B7C CRC64;

Query Match 38.8%; Score 718; DB 5; Length 218;
Best Local Similarity 59.2%; Pred. No. 2.6e-51;
Matches 129; Conservative 37; Mismatches 52; Indels 0; Gaps 0;

OY 1 MSPILGWKIKGLVPTRLLEYLEEKYEHLHYERDEGDKMKNKFELEFPNLPYYID 60
DB 1 MAPVLGYWKIRGLADPIRLLELYGDSYEHSHYGRKDGKMDKHNGLLEPPLPYKD 60
OY 61 GGVKLTQSMALIRYIADKHNMLGGCPKERAETSMLEGAVALDIRYGVSRIVASKDEFETLKV 120
DB 61 GNFSLTQSLAIRYIADKHNMLGMPVERAKISMIEGGLVDIRAGVSRIVAOETFEOLKV 120
OY 121 DFLSKLPKMFEDRLCKHTYLNCDHVTHPDFMLYDALDVLVYMDPMLDAPFKLVCFK 180
DB 121 DFLNLPKMFEDRLCKHTYLNCDHVTHPDFMLYDALDVLVYMDRCKLDAFPKLVCFK 180
OY 181 KRIEAIPOIDKYLKSSKYIAMPLOGQATFGGDDHPK 218
DB 181 KRIENIPOINEYLKSSKYIEMPLQWQATFGGDDHPK 218

RESULT 8
OY 09XYL9 PRELIMINARY; PRT: 218 AA.
AC 09XYL9;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, last annotation update)
DE GLUTATHIONE S-TRANSFERASE.
GN GST-1.
OS Fasciola gigantica.
OC Eukaryota; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea;

```

```

OC Echinostomida; Echinostomata; Fascioloidae; Fasciolidae; Fasciola.
OX NCBI_TaxID=46835;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-THAILAND.
RA Grams S.V., Grams R., Sobhon P., Vivanant V., Upatham E.S.;
RT "Molecular cloning of expressed antigens from Fasciola gigantica."
RL Submitted (DEC-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF112567; AAD23997.1;
DR HSSP: P31670; 1FHE.
DR INTERPRO: IPR000521;
DR PFM: PFO0043; GST; 1.
KM Transferase.
SQ SEQUENCE 218 AA; 25337 MW; E9423D75C3F2EEAF CRC64;

Query Match 37.7%; Score 698; DB 5; Length 218;
Best Local Similarity 56.7%; Pred. No. 1.1e-45;
Matches 125; Conservative 35; Mismatches 53; Indels 0; Gaps 0;

OY 5 LGYWKIKGLVPTRLLEYLEEKYEHLHYERDEGDKMKNKFELEFPNLPYYIDGVK 64
DB 5 LGYWKIRGLQDPVRLLEYLEDEYEHLYXGRDKRMKLGDFNMGDLDPNLPYYIDCKK 64
OY 65 LTQSMALIRYIADKHNMLGGCPKERAETSMLEGAVALDIRYGVSRIVASKDEFETLKVDFLS 124
DB 65 LTQSAIWMRYIADKHNMLGSTPEERARVSMIEGAMDLRMGEFVRCVYNPNEEYVGDYK 124
OY 125 KLPEMLKMFEDRLCKHTYLNCDHVTHPDFMLYDALDVLVYMDPMLDAPFKLVCFK 184
DB 125 ELPKTKMWSDFLGRQYLTGSSVSHVDFWYEAIDCIRYLAPOCLNDFPKLKRSHIE 184
OY 185 AIPQIDKYLKSSKYIAMPLOGQATFGGDDHPK 217
DB 185 DLPKIKAYMESKPIKMPPLNLSWTASFSGGDAAP 217

RESULT 9
OY 09NOV4 PRELIMINARY; PRT: 218 AA.
AC 09NOV4;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, last annotation update)
DE CLASS MU GLUTATHIONE S-TRANSFERASE.
GN GSTM.
OS Bos taurus (bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Tissue-eye lens;
RA Jimenez-Jasento J.V., Garland D.;
RT "A lens glutathione S-transferase, class mu, with thiol-specific
RT antioxidant activity."
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF249588; AAF64308.1;
KM Transferase.
SQ SEQUENCE 218 AA; 25635 MW; 30D2EA0F43C07B0A CRC64;

Query Match 27.7%; Score 512.5; DB 6; Length 218;
Best Local Similarity 47.6%; Pred. No. 1.8e-34;
Matches 99; Conservative 35; Mismatches 69; Indels 5; Gaps 1;

OY 1 MSPILGWKIKGLVPTRLLEYLEEKYEHLHYERDEG----DKRNKKFELGEPNVL 55
DB 1 MPMILGYWDIRGLAIAIRLLELYTDNYEEROYSVGADAPDYDRSOWLMEKFGIDPFLN 60
OY 56 PYIINGDVKLTQSMALIRYIADKHNMLGGCPKERAETSMLEGAVALDIRYGVSRIVSKDEF 115
DB 56 PYIINGDVKLTQSMALIRYIADKHNMLGGCPKERAETSMLEGAVALDIRYGVSRIVSKDEF 115

```


OC Echinococcus granulosus.
 OC Eukaryota: Metazoa: Platyhelminthes: Neodermata: Cestoda; Eucestoda.
 OC Cyclophyllidae, Taeniidae; Echinococcus.
 OX NCBI_TaxID=6210;
 RN (1)
 RN
 RP SEQUENCE FROM N.A.
 RP Fernandez V., Zaha A., Fernandez C.;
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Fernandez Mancebo V., Chalar C., Martinez C., Zaha A.,
 RA Fernandez Granja C.;
 RT "EGGS": a glutathione S-transferase gene from Echinococcus
 RT granulosus";
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF005928; AB66318.1; -;
 DR EMBL; AF101269; AB16438.1; -;
 DR HSSP; P20136; 1GSU.
 DR INTERPRO: IPR000521; -;
 DR PFAM: PF00043; GSTR, 1.
 KW Transferase.
 SO SEQUENCE 219 AA; 2553 MW; 8883E70AD075D154 CRC64;

```
. Query Match          25.6% Score 473.5; DB 5; length 219;
   Best Local Similarity 42.4%; Pred. No.2.9e-31;
Matches    92; Conservative    40; Mismatches    80; Indels     5; Gaps      1.
```

Search completed: March 18, 2001, 05:37:15
Job time: 463 sec

